Int. J. Human–Computer Studies 179 (2023) 103106

Contents lists available at ScienceDirect

International Journal of Human - Computer Studies

journal homepage: www.elsevier.com/locate/ijhcs

Speaking in front of cartoon avatars: A behavioral and psychophysiological study on how audience design impacts on public speaking anxiety in virtual environments

Matteo Girondini[a,b,*], Milena Stefanova[a,c], Margherita Pillan[c], Alberto Gallace[a,b]

[a] Mind and Behavior Technological Center, Department of Psychology, Università Milano-Bicocca, Italy
[b] Department of Psychology, Università Milano-Bicocca, Italy
[c] Department of Design, Politecnico di Milano

ARTICLE INFO

Keywords:
Virtual reality
Multisensory perception
Public speaking anxiety
Human-computer interactions
Sense of presence

ABSTRACT

Public speaking anxiety is defined as a strong difficulty in speaking in front of an audience and has been shown to impair work performance and social relationships. Virtual Reality (VR) offers an efficient tool for modulating public speaking anxiety through a wide range of customizations concerning environmental settings. However, scientific research needs to understand better what features of the simulated environment are more important to increase or reduce participants' perceived discomfort. The present study investigates the role of visual (human vs cartoon characters) and acoustic (human vs robotic voice) audience features on perceived anxiety, sense of presence, and perceived realism in an interactive VR public speaking scenario. *42* participants *(mean age = 24 y. o; Females = 30)* performed four public speaking sessions, characterized by different levels (high vs. low) of graphic and acoustic audience design. Both explicit (questionnaires) and implicit physiological measures (Electrodermal activity-EDA) collected during audience interaction were used to assess the participants' experiences.

The results showed that the features of the simulated audience played a crucial role in perceived anxiety during a virtual public speaking. Specifically, the more realistic level of graphic and acoustic stimuli resulted in higher levels of self-reported anxiety as compared to the lower realism level. However, the experienced realism and the sense of presence seem more affected by the graphical than acoustic features of the virtual environment. By contrast, the acoustic features impact on the interaction realism with the virtual audience. Interestingly, the robotic voice (lower acoustic realism) increased electrodermal response during the interaction with the audience, interpreted as a break in the sense of presence. A positive correlation between anxiety, sense of presence, and experienced realism was found. As well, perceived anxiety is correlated with electrodermal activity during the performance. Nevertheless, physiological activity is more affected by the first experience than the realism features, suggesting the presence of a habituation effect across the repeated sessions. Taken together, the results of our study showed that multisensory features (graphical and acoustic) of the virtual environment play a fundamental role in creating realistic public speaking experiences and might be used within gamification strategies for soft skill training (e.g., for improving public speaking anxiety).

1. Introduction

1.1. Public speaking anxiety

Public Speaking Anxiety (PSA) has been described as a sense of discomfort and distress to perform a discussion or speech in front of an audience (Pull, 2012). This condition is quite common during social interactions, but in some cases, the anxiety evoked by public speaking becomes pathological (Ferreira Marinho et al., 2017; Sandra R Harris et al., 2002; Heimberg, 2002; Hoffman et al., 2004; Raja, 2017; Stein and Chavira, 1998). Regardless of its clinical manifestation, public speaking can be perceived as an unpleasant and threatening experience, characterized by high stress levels and bodily discomfort, even in non-clinical populations. Further evidence of the effectiveness of public

* Corresponding author at: Department of Psychology, University of Milano-Bicocca, Piazza dell'AteneoNuovo, 1, 20126 Milan, Italy.
E-mail address: m.girondini@campus.unimib.it (M. Girondini).

https://doi.org/10.1016/j.ijhcs.2023.103106
Received 12 December 2022; Received in revised form 28 June 2023; Accepted 4 July 2023
Available online 11 July 2023

speaking as a stressful event is also suggested by the inclusion of this task into the *Trier Social Stress Test*, one of the most widely used experimental settings for inducing acute psychosocial stress under laboratory conditions (Birkett, 2011; Fich et al., 2014). PSA, even when nonpathological, can have a substantial impact on work opportunities and social relationships, deteriorating the well-being and career choices of those who are affected by it (Aderka et al., 2012; Hinojo-Lucena et al., 2020).

Public speaking anxiety is characterized by behavioral symptoms and an increase in physiological arousal before and during speaking. Behavioral outcomes of PSA concern the avoidance of performance, the anticipation of anxiety, embarrassment, and a state of discomfort, altogether increasing the risk of making communicative mistakes (Klinger et al., 2005; Palmas et al., 2021; Panayiotou et al., 2017a; Stein and Chavira, 1998). Additionally, public speaking anxiety is recognized to trigger a significant increase in physiological arousal and somatic tension before, during, and after the task (Goodman et al., 2017). This effect can be described in terms of the sympathetic nervous system activation in response to a real or imagined/anticipated threat (Behnke and Sawyer, 2001; Croft et al., 2004; Finn et al., 2009; Giesen and McGlynn, 1977; Pörhölä, 2002). Concerning this aspect, several parameters are often used to assess the physiological response during public speaking task, such as Heart Rate (HR) (Bassett et al., 1987; Lacy et al., 1995; Panayiotou et al., 2017b; Panknin et al., 2002; Premkumar et al., 2021), Heart Rate Variability (HRV) (Åhs et al., 2009; Kothgassner et al., 2016; MacIntyre et al., 2010), Electrodermal activity (Skin conductance level -SCL, or Skin conductance response -SCR) (Carrillo et al., 2001a; Dimberg et al., 1986; Giesen and McGlynn, 1977; Gonzalez-Bono et al., 2002; McKinney and Gatchel, 1982; Panayiotou et al., 2017a; Parente et al., 2005). These parameters, together with self-report measurements, provide an additional indicator in respect with the anxiety experienced by participant (Croft et al., 2004; Schwerdtfeger, 2004).

The possibility of analyzing physiological parameters during a task is also relevant to the question regarding the relationship between anxiety and performance. That is, the well-known Yerkes-Dodson law (Chaby et al., 2015; Green et al., 1908; Teigen, 1994) suggests that increases in physiological arousal up to a certain level lead to better performance, whereas further increases beyond that point deteriorate people's abilities to execute the task (following a bell-shaped function relating performance and physiological arousal). Thus, physiological parameters may be taken as a further predictor of human performance in several challenging situations, such as public speaking.

1.2. Public speaking in virtual reality (VR)

Nowadays, Virtual Reality (VR) is a promising tool for developing exposure therapies and, more in general, for skill training through vivid sensory-motor experiences Campo (Bush, 2008; North et al., 1998; Powers and (Bush, 2008; North et al., 1998; Powers and Emmelkamp, 2008). This occurs because, if correctly designed, VR illusions (i.e., the perception of acting within an artificial environment just like in a real one) can activate the same neurological mechanisms that allow participants to respond to real features of the environment (Gallace et al., 2011; Higuera-Trujillo et al., 2017; Lanier, 2017; Riva, 2009). Immersivity plays a key role in creating a reliable experience inside a virtual environment, which is measurable through the subjective 'sense of presence'. The sense of presence refers to "being there" in a simulated environment (Slater, 2002; Slater et al., 2009b; Slater and Wilbur, 1997). Presence in virtual environments is usually measured by means of self-report questionnaires. Still, more recently, a number of physiological measurements have been used as "markers" regarding the degree of presence experienced by the participant (Bystrom et al., 1999; Grassini and Laumann, 2020; Insko, 2003).

In the last twenty years, several studies reported that Public Speaking tasks performed in VR are capable of evoking physiological activation and psychological discomfort (Owens and Beidel, 2015; Pertaub et al., 2001; Premkumar et al., 2021; Reeves et al., 2021; Šalkevičius et al., 2019; Slater et al., 2006b; Stupar-Rutenfrans et al., 2017), comparable to their real counterparts (Ebrahimi et al., 2019). The effectiveness of VR applications regarding public speaking anxiety is also proven by the different responses evoked by the simulated environment (in terms of behavioral and physiological activation) by high vs. low trait anxiety participants during the experimental manipulations (Gonzalez-Bono et al., 2002; Palmas et al., 2021). Not only VR is a good tool for clinical treatment (Bouchard et al., 2017; Carl et al., 2019; Sandra R. Harris et al., 2002; Kahlon et al., 2019; Kampmann et al., 2016; Premkumar et al., 2021; Reeves et al., 2021), but it may be considered a suitable tool for human soft and hard skill activity training (e.g., public speaking; work interviews, etc.) in non-clinical populations (Takac et al., 2019). In fact, through realistic and immersive VR experiences, participants can perform public speaking activities in a controllable and safe setting. They can be supported against personal distress, arousal activation, and distorted cognitive perception during the exposure (Kahlon et al., 2019). Each session can also be used to manipulate the main focus of the exercises (i.e.: speech circumlocution, body language, eye contact with the audience) and thus be customized to the client's needs (Šalkevičius et al., 2019). The environmental setting (e.g., room characteristics, size of the audience, etc.) can be programmed as a function of the specific user or company requirements (i.e.: work discussions, slide presentations, job interviews, oral school examinations, reports to the company board, etc.). Public speaking in VR is also open to interactions with other technological advancements. For example, recent development in the field of AI (Artificial Intelligence) will be used in the near future to enhance the realism of the simulated interaction by modifying certain features of the scenario (such as overall level of noise; patterns of light changes) or of the autonomous agents' reactions (i.e. verbal, motor and visual responses) on the basis of the actual actions, physiological responses, and behaviors of the speaker (El-Yamri et al., 2019). It should be noted, however, that for this possibility to be fully implemented, the system must be trained to analyze certain behaviors and physiological parameters from the users and select those that are relevant to the given situation to be simulated (e.g., imagine a speaking scenario that produces more hostile responses from the listeners when the overall level of users' arousal, measured through increases in heartbeat or sweating, increases). This certainly requires a better knowledge of the participants' responses to different VR environmental features.

1.3. The role of realism in VR public speaking design

Despite the importance of contextual variables in the VR environment in affecting participants' responses (i.e., audience behaviors (Chollet et al., 2018; Pertaub et al., 2002; Premkumar et al., 2021), environment architectural design (Fich et al., 2014), or audience behavior (Girondini et al., 2023)) certain parameters related to the design of the virtual scenario where public speaking occurs have been very little explored. It has been hypothesized that visual realism (here defined as the perceived similarity between a real scene and a simulated one) might enhance realistic behavioral responses (Slater et al., 2009a; Vinayagamoorthy et al., 2004), but not much is known about how the features of a simulated audience (e.g., human-like vs. cartoon-like avatars) can impact a person's anxiety and sense of presence (Goncalves et al., 2021). This is an important aspect of VR research for two reasons: 1) overly realistic avatars might lead to the well-studied Uncanny Valley effect (Geller, 2008; Ho and MacDorman, 2017; Mori et al., 2012; Tinwell, 2014); 2) cartoon-like characters might offer many more possibilities than real human avatars to be modulated in terms of alternative bodily features (e.g., in order to be perceived as less threatening, for example by increasing the size of the eyes or the roundness of the body; (Partala and Surakka, 2003)) and might be used within some gamification strategies (Pereira et al., 2014). As far as gamification (application of game elements in a non-gaming context) is concerned, this strategy has already been used to improve human learning processes, with positive outcomes (Grivokostopoulou et al., 2017; Pereira et al., 2014; Ryu et al., 2018; Sailer and Homner, 2020; Su, 2016). The game design applied to a stressful experience, such as public speaking, could help to distract the user from the primary task but at the same time maintain the learning and habituation benefits of the exposure training (Yee, 2006). Using a cartoon-like avatar within a gamification strategy might thus help to modulate anxiety about one's own performance and improve self-efficacy before approaching a more stressful real human audience. Recently, Bellido Rivas and colleagues (Bellido Rivas et al., 2021) have compared anxiety perceived after a VR public speaking training performed in a cartoon scenario (with the speaker embodied in a cartoon) or within a more realistic (human-like avatars) scenario. The study did not confirm the advantage of using cartoon-like avatars on perceived anxiety, so other clarifications are required.

1.4. Multisensory interaction in virtual environments

In order to investigate the impact of audience features within virtual public speaking, a multisensory perspective might be useful. In fact, like real interactions, most virtual interactions are based on multisensory integration processes occurring inside the brain (Gallace et al., 2011). The participants' sense of presence and perception of realistic experience inside the virtual environment could be enhanced by integrating coherent sensory signals (Gallace et al., 2011; Marucci et al., 2021). In this instance, a visually realistic scenario might lose its level of effectiveness whenever acoustic details of the interaction are not realistic (or incongruent from a multisensory point of view) (Frassinetti et al., 2002; Maffei et al., 2016). Matching visual and acoustic elements is relevant for inducing a realistic virtual interaction (Mitchell et al., 2011). On the other hand, different combinations of visual and acoustic information (more or less coherent) can lead to different levels of public speaking anxiety, and studying their role and interaction can help to create better exposure therapies or skill development training procedures. For this very reason, it becomes relevant to analyze both the visual and acoustic elements of the virtual environment, especially during the interaction between participants and the simulated audiences during the virtual session (such as an interactive dialog), and how this may be perceived as "realistic". Naas has claimed that humans can be socially influenced by virtual avatars just as they would be in the real world (Nass et al., 1994) and this element is becoming even more important nowadays with the technical feasibility of reproducing 'metaverse'- based virtual interactions (virtual spaces shared between multiple users; (Dionisio et al., 2013; Gallace, 2022; Wiederhold and Riva, 2022). Not surprisingly, virtual reality has been used to reproduce socially relevant contexts, such as 'social dilemma' paradigms or emotionally relevant interactions (Pan et al., 2011; Skulmowski et al., 2014; Volante et al., 2016). This aspect also leads to another important theoretical concept that might be relevant in the case of virtual interactions between multiple players (some or all of them being human-controlled, and some being AI-controlled) known as "social presence" (Biocca et al., 2003; Guimarães et al., 2020; Oh et al., 2018a; Wakeford, 2002). Research on simulated virtual interactions needs to define which proprieties of the environment can evoke realistic responses by the user during a simulated social interaction (between the user and the virtual characters). Even though this question arose more than twenty years ago, the conditions necessary to elicit a realistic human-avatar interaction remain rather unclear (Kyrlitsias and Michael-Grigoriou, 2022; Wienrich et al., 2018). Earlier studies on social presence reported no difference when the participant interacted with virtual humans or with virtual avatars that did not look human (Nowak and Biocca, 2003; Soash, 1999; Volante et al., 2016). Despite this consideration, in the last decade, a significant improvement in the graphical rendering of virtual experience has occurred, and some studies failed to confirm the previous findings (Dubosc et al., 2021; Kwon et al., 2013; Yoon et al., 2019). More interestingly, some works suggested that other factors, such as audio and/or animation features, are equally relevant than graphical rendering within virtual interaction in computer-generated environments (Chollet and Scherer, 2017; de Melo et al., 2013; McDonnell et al., 2012; Volante et al., 2016). Among them, behavioral realism (the level of similarity between virtual behavior by the avatar and real-world behavior of humans in a given situation) is a further element that is able to increase the sense of social presence and realism of virtual interaction (Herrera et al., 2020). Virtual social interaction involving avatars characterized by high behavioral realism might enhance social presence and influence users' behavior (Guadagno et al., 2007). For this reason, behavioral realism must be considered (and/or controlled) while evaluating how graphical and acoustic realism impact the simulation experience (Garau, 2003). Furthermore, the design of virtual social interactions is particularly relevant in the case of public speaking since realistic interplay and its modulation as a function of certain environmental features might be task-related (Allmendinger, 2010). That is, it is reasonable to hypothesize that very realistic multisensory interactions in high-pressure or stressful situations, such as public speaking, could result in greater discomfort, particularly for high-anxiety users (Palmas et al., 2021). By contrast, lower perceived realism (based on the visual or acoustic properties of the avatars) in the same situation might reduce or mitigate the perceived anxiety during the task. In summary, within public speaking scenarios, the visual-acoustic features of the avatar might play an important role. Modulating the multisensory details of avatars might help us understand how public speaking anxiety can be affected by the design of the virtual environments in which they take place.

1.5. Current study

1.5.1. Aim of the study and research hypotheses

The main aim of the current study is to test the effect of audience graphical-acoustic features, by means of different VR interactive scenarios, on participants' anxiety levels and sense of presence during public speaking experiences. The study also aims to understand whether our modulations differently affect more explicit (e.g., anxiety self-assessment) or implicit (physiological responses) measures in a non-clinical population. Given the possible differences driven by trait anxiety (Gonzalez-Bono et al., 2002; Palmas et al., 2021), before starting the VR session, a wide range of anxiety self-assessment scales (concerning classic anxiety questionnaires and social anxiety questionnaires) were included. The experiment was composed by four different Public Speaking VR scenarios, and each scenario differed from the others in terms of visual and acoustic information presented. In order to investigate the effect of audience design in respect to virtual interactions, all of the scenarios contained an interactive session characterized by a question asked by the audience.

Based on the important role played by multisensory interactions in people's perception, also confirmed within virtual environments, it is reasonable to hypothesize:

H1. A virtual audience with a high level of graphical realism, compared to one with low graphical realism, elicits more anxiety during public speaking tasks

H2. A virtual audience with a high level of acoustic realism, compared to one with low acoustic realism, elicits more anxiety during public speaking tasks

H3. The acoustic and graphical realism of the audience may have a two-way interaction effect on anxiety during public speaking tasks, which indicates that the effect of graphical realism on public speaking anxiety is different at different levels of acoustic realism, and viceversa.

Secondary analyses concerned the impact of graphical and acoustic realism on the sense of presence, the sense of realism, and interaction realism. Moreover, we decided to analyze also the physiological response to the audience question and during the whole performance, indexed by the phasic activity of electrodermal activity (SCR, NS-SCR). A correlation analysis was performed between perceived anxiety, subjective reports, and physiological activity.

2. Methods

2.1. Materials

2.1.1. Sample size

The participants were recruited by self-enrollment using the University web site. University students were given course credit for participating. The sample size for this study was calculated using G*Power 3.1 (Faul et al., 2007) and the repeated ANOVA statistical test (four measurements and two within factors). G*Power indicated that, with $\alpha = 0.05$ power (1-β) = 0.80, and a medium effect size (0.4), the estimated sample size for this study was at least 42 participants.

A total of *N = 42* participants aged between 20 and 63 years comprised the experimental sample (*M = 24, SD = 13.2)*. There were twenty-nine females *(70%)* and thirteen males *(30%)*. Thirty-three participants admitted having previous experiences with Virtual Reality *(79%)*, while nine *(21%)* had not. The study was conducted according to the ethical principles of the Declaration of Helsinki and was approved by the Department of Psychology, University of Milano Bicocca, local ethical committee.

2.1.2. Experimental design

The experimental design involved participants performing four public speaking performances, each with varying levels of graphical and acoustic realism. Specifically, the audience could be characterized by high (human avatar) vs. low (cartoon avatar) ***graphical realism*** and high (human voice) vs. low (robotic voice) ***acoustic realism,*** resulting in a 2 × 2 repeated measure design. The scenario was either high or low in both aspects, and the order of presentation was counterbalanced across participants. This resulted in the following four experimental scenarios:

- **High Visual Realism – Human Voice (H_H):** The scenario was composed of a human avatar audience and the question was presented by using a human voice.
- **High Visual Realism – Robotic Voice (H_R):** The scenario was composed of a human avatar audience, but the question was presented by using a robotic voice
- **Low Visual Realism – Human Voice (L_H):** The scenario was composed of a cartoon-like audience and the question was presented by using a human voice
- **Low Visual Realism – Robotic Voice (L_R):** The scenario was composed of a cartoon-like audience and the question was presented by using a robotic voice

The primary outcome measure was the participants' perceived anxiety after each performance. Secondary analyses included subjective reports of sense of presence, interaction realism, and perceived realism comparing high vs. low graphical and acoustic realism. Additionally, electrodermal activity during all the performances was recorded and analyzed. Finally, correlation analyses were conducted between subjective reports, anxiety, and physiological activity. Overall, this design served us to investigate how different levels of graphical and acoustic realism impact anxiety and physiological responses during public speaking performances in virtual reality.

2.1.3. VR equipment and scenarios

The VR equipment used for the experiment included an Oculus Rift Head mounted display (HMD), with a 1280 × 1440-pixel resolution per eye (refresh rate 80 Hz). The HMD was connected to a computer, featuring an Intel Core i7–7800X CPU, 16 GB of RAM and a GeForce GTX GPU.

Two VR environments were developed with the Unity graphics Engine (https://unity.com/.) for the purposes of this experiment: – high (Fig. 1) and low (Fig. 2) visual realism audience. For each of them, two audio versions were prepared on the basis of the voice added to the audience characters, resulting in high (human voice) vs (low robotic voice) acoustic realism. All scenarios involved a room where eleven chairs were arranged. Eight avatars were sitting in some of these chairs. A virtual projection board hanging on the wall just behind the audience and in front of the speaker showed the sequence of the different pictures involved in the preparation of the recipe being presented by the participant. Adobe Fuse was used for the creation of the humanlike avatars. Body movements were selected from the Adobe Mixamo library (https://www.mixamo.com) and were then combined and blended into Unity with the use of layers and masks. All members of the audience are in a sitting position. Movements include turning head left or right, head nodding, scratching a hand and moving on the chair. The gestures are neutral and do not represent a positive or negative reaction from the people in the audience. For the question session, mouth movement was implemented in Unity and added to one of the audience avatars. The avatars in the low visual realism environment were created with Blender graphics software (https://www.blender.org/). Their shape was vaguely reminiscent of Barbapapa cartoon characters (https://www.barbapapa.com/the-barbapa-family-en/), with rounded and colored bodies, which included eyes, eyebrows, and a mouth. All animated body movements were created with the Unity graphics engine (i.e., Nodding in agreement, Looking the presenter in the eyes). The movements from the humanlike scenario are translated into a similar movement for the cartoon avatars. The timeline and the number of gestures is kept the same between the two visual settings. During each exposure, after 3 min of participants' speech, one element of the audience (human avatar or cartoon avatar) raised his/her hand and asked a question (with human or robotic voice) related to the speech topic. The voice for each question used was recorded offline by people naïve to the experiment and inserted in Unity. In the case of robotic voice, a human voice was modified by a robot voice generator app (*Magicmic robotic voice changer)*. The questions were similar in terms of length (around 8–10 s.).

2.2. Measurements

2.2.1. Anxiety questionnaires

STAI (State-Trait Anxiety Inventory): STAI is a well-known anxiety self-assessment scale, composed of two subscales which measure transient and enduring levels of anxiety (Beckler, 2010; Spielberger et al., 1971). Each scale has 20 items, and both scales include items that describe symptoms of anxiety and items that describe the absence of anxiety. We used the Italian language validated version of STAI.

SIAS (Social Interaction Anxiety Scale): SIAS is a self-report questionnaire that measures the presence of fear during general social interactions. The Italian translation contains 19 items evaluated on a 5-

Fig. 1. Scenario for high graphical realism.

Fig. 2. Scenario for low graphical realism.

point Likert scale, rating from 0 (not at all) to 4 (extremely) (Heimberg et al., 1992).

*SPS (Social Phobia Scale):*SPS is a self-report questionnaire that measures anxiety in situations involving being observed by others. The Italian translation contains 20 items evaluated on a 5-point Likert scale, rating from 0 (not at all) to 4 (extremely) (Heimberg et al., 1992).

SSQ (Simulator Sickness Questionnaire): SSQ is a motion sickness questionnaire used to assess sickness after virtual reality experiences. The version used in this experiment contains 16 items, divided into three categories: nausea, oculomotor problems and disorientation (Kennedy et al., 2009).

2.2.2. *Experience evaluation (self-report)*

Four different self-report statements were used to assess participants' subjective experiences during the task. Each statement was evaluated by using a 10-point Likert scale, ranging from 0 (not at all) to 10 (extremely). The measurements were collected after each experimental session. The focus of Likert assessments concerned:

- *Sense of presence* (I felt like I was inside the environment shown)
- *Experienced realism* (The content of this environment seemed credible to me)
- *(Audience) Interaction realism* (The interaction I had with the avatar seemed believable and realistic)
- *Level of anxiety evoked by the task* (How much anxiety did this experience evoke from 1 to 10)

2.2.3. *Physiological measurement*

Electrodermal activity (EDA): The measurement of the phasic level of skin conductance is recognized as a highly suitable marker for sympathetic nervous system activation (Turpin and Grandfield, 2007). Other studies already used EDA as a marker of anxiety state during public speaking (Arsalan and Majid, 2021; Giesen and McGlynn, 1977; Westenberg et al., 2009). For example, Croft and colleagues used EDA values to predict state-dependent speech anxiety in a student sample during a public speaking task (Croft et al., 2004). The focus of EDA measurement during public speaking was related to the phasic (fast) change of electrodermal activity. In particular:

- Skin Conductance Response (SCR): Changes of the phasic level of Skin conductance can be found during stressful situations and have been shown to reliably indicate the presence of mental stress (Lawler, 1980), challenging situations, or unexpected environmental conditions (Connolly and Frith, 1978; Harris and Katkin, 1975; Zeiner and Smith, 1979). Dimberg and colleagues used SCR to compare social stimuli to neutral ones in different participants during public speaking tasks (Dimberg et al., 1986).
- Non-specific Skin Conductance Response (NS-SCR): The NS-SCR is the frequency of phasic level of electrodermal activity that occurs spontaneously, not related to external stimuli (in a fixed time-window) (Carrillo et al., 2001b; Gertler et al., 2020; Nikula, 1991). This measure has been previously used as a measure of fear-induced arousal during public speaking situations (Niles et al., 2015).

For Skin Conductance measurements, in the present study, a Biopac BioNomadix MP 150 device recorded the electrodermal signal by means of two AgCl electrodes attached to the participant's index finger and ring finger. The signal was recorded at 100 Hz and downsampled to 10 Hz for the analysis. Pre-processing of the raw data was performed using Acqknowledge software (www.biopac.com). To minimize the presence of artifacts in the data, a 1 Hz low-pass filter was applied. We extracted two measures of interest for each exposure, which lasted approximately 4.10 min: the number of non-specific skin conductance responses (NS-SCR) and the skin conductance response amplitude (SCR). For each scenario, we extracted the number of NS-SCR and used it as the dependent variable in the analyses. To compute SCR, a threshold of 0.03 Microsiemens was applied. For each event (i.e., the onset of the question by the virtual audience), we collected data within a 6–8 second time window, as recommended by prior research (Braithwaite et al., 2013). The amplitude of the SCR was extracted for each question and used as the dependent variable.

2.3. *Procedures*

Each participant completed all four scenarios during the experimental session, with a counterbalanced order (in terms of which scenario was presented as the first scenario) across participants. The task for the participant was to explain a cooking recipe (studied for 5 min, just before the exposure) to the virtual audience. The recipe was changed in each of the four scenarios (the chosen recipes were comparable in terms of difficulty and length). Moreover, the participant was not informed of the interaction with the audience before the first speaking session.

The participants arrived at the laboratory and signed the informed consent form. Demographic and anxiety self-assessment questionnaires were provided by means of a tablet or computer. Upon completion of all questionnaires, the participants received a sheet of paper containing the first cooking recipe to study, and they were given up to 6 min to memorize all of the information. They were also told that their task was to illustrate the entire recipe to an audience. They were informed that recipe images would appear behind the audience during the task (like speech-related slides). When ready, the participants were taken to the VR station for the recording of a physiological baseline (two minutes). After the baseline was recorded, they were asked to wear the Head-Mounted-Display and the simulation started, together with the recording of physiological data (Fig. 3). After four minutes, the simulation was stopped, and the participants completed the Likert assessment. At the end of the first experimental condition, they received the second cooking recipe under the same experimental procedures. The timeline was the same for all four experimental conditions. Once the VR experience was concluded, the final step was the assessment of VR sickness using the SSQ questionnaire (Fig. 4).

2.4. *Data analysis*

Statistical analyses were performed by R software (www.r-project.org) and reported in the OSF repository (https://osf.io/d37e6/files/osfstorage). To explore the relationship between perceived anxiety and subjective reports of the VR experience, correlation matrices using the Pearson correlation coefficient (r) were computed. Specifically, we calculated the correlations between perceived anxiety and two measures of subjective experience: the sense of presence and experienced realism. Additionally, we conducted a correlation analysis to examine the

Fig. 3. Participant during VR public speaking exposure.

associations between the sense of presence, self-reported anxiety, and physiological arousal measured by the number of non-specific skin conductance responses (NS-SCR).

A repeated measures ANOVA was conducted on perceived anxiety, subjective reports (i., e., sense of presence, experienced realism, and interaction realism), and physiological data (SCR, NS-SCR) with two within-subject factors: *graphical realism* (high vs. low) and *acoustic realism* (high vs. low). The ANOVA and regression analysis showed robustness against violation of normality distribution due to using the Likert scale (Norman, 2010), see also, Higgins et al., 2022). In case of sphericity violations, Greenhouse-Geisser correction was applied. A pairwise comparison with Bonferroni correction was conducted where significant interaction effects were reported. The results included p-values (set as < 0.05) and effect size (generalized eta square).

3. Results

3.1. Descriptive results: anxiety scale and VR sickness

The average anxiety score revealed normal values, confirming that our participants belonged to a nonpathological sample: STAI-S (*$M = 34.5$, $SD = 7.28$*), STAI-T (*$M = 44.1$, $SD = 7.75$*), SIAS (*$M = 24.5$, $SD = 8.31$*) and SPS (*$M = 20$, $SD = 10.1$*). In order to detect the possible presence of some form of sickness due to the device used, we analyzed the data from SSQ questionnaires completed after the VR exposure. The SSQ scores revealed a low level of sickness (*$M = 3.13$, $SD = 2.68$*), suggesting that the participants well tolerated our VR public speaking experiences.

3.2. Self-report measurements

Sense of Presence: A significant effect of *graphical realism ($F = 33.61$, $df = 1,41$, $p < .001$, $\omega2 = 0.063$)* regarding interaction realism during public speaking was found. Participants reported a higher sense of presence after the avatar audience (high graphical realism, $M = 6.66$, $SD = 1.99$) than the cartoon audience (low graphical realism, $M = 5.47$, $SD = 2.59$). No main effect of *acoustic realism ($F = 0.48$, $df = 1, 41$, $p = 490$, $\omega2 < 0.001$)* or interaction effect *graphical realism * acoustic realism ($F = 0.59$, $df = 1, 41$, $p = .445$, $\omega2 < 0.001$)* was found (Fig. 5).

Experienced realism: A significant effect of *graphical realism* (*$F = 86.42$, $df = 1,41$, $p < .001$, $\omega2 = 0.28$*) concerning experienced realism during the exposure was found. Participants reported higher experienced realism after the avatar audience (high graphical realism, $M = 7.01$, $SD = 1.93$) compared to the cartoon audience (low graphical realism, $M = 4.31$, $SD = 2.39$). No main effect of *acoustic realism ($F = 0.73$, $df = 1, 41$, $p = .397$, $\omega2 < 0.001$)* or interaction effect *graphical realism * acoustic* realism was found *($F = 0.02$, $df = 1, 41$, $p = .887$, $\omega2 < 0.001$)* (Fig. 6).

(Audience) Interaction realism: A significant effect of *graphical realism ($F = 37.68$, $df = 1,41$, $p < .001$, $\omega2 = 0.187$)* regarding interaction realism during public speaking was found. Participants reported higher interaction realism after the avatar audience (high graphical realism, $M = 6.67$, $SD = 1.90$) than the cartoon audience (low graphical realism, $M = 4.60$, $SD = 2.46$). A main effect of *acoustic realism ($F = 7.29$, $df = 1, 41$, $p < .001$, $\omega2 = 0.020$)* was also found, indicating that the human voice (high acoustic realism, $M = 5.95$, $SD = 2.37$) induced a higher sense of interaction realism compared to the robotic voice (low acoustic realism, $M = 5.33$, $SD = 2.45$). A non-significant trend (*$F = 3.87$, $df = 1,41$, $p = .055$, $\omega2 = 0.004$*) for the interaction of *graphical realism * acoustic realism* was present, suggesting that the difference between human avatar vs. cartoon avatar, in terms of interaction realism perceived, was higher in case of human voice compared to the robot voice (Fig. 7).

Level of anxiety evoked by the task: A main effect of *graphical realism ($F = 22.86$, $df = 1, 41$ $p < .001$, $\omega2 = 0.036$)* was found. Participants reported higher anxiety after the avatar audience (high graphical realism, $M = 4.98$, $SD = 2.15$) compared to the cartoon audience (low graphical realism, $M = 4.14$, $SD = 2.20$) (Fig. 9). A main effect of *acoustic realism ($F = 8.03$, $df = 1, 41$ $p < .001$, $\omega2 = 0.018$)* was also found, indicating that the human voice (high acoustic realism, $M = 4.85$, $SD = 2.32$) induced higher anxiety compared to the robotic voice (high acoustic realism, $M = 4.27$, $SD = 2.07$). No interaction effect *graphical realism * acoustic realism* was found *($F = 0.98$, $df = 1, 41$ $p < .325$, $\omega2 = 0.001$) (Fig. 8).*

3.3. Physiological results

Regarding the physiological results, three participants were not included in the analysis due to artifacts in the signals ($n = 1$) and data storage problems ($n = 1$). Thus, the physiological analysis included a total sample of 40 participants.

Non-specific skin conductance response (NS-SCR): No main effect of *graphical realism ($F = 0.49$, $df = 1,39$, $p = .486$, $R^2 < 0.001$), acoustic realism* ($F = 1.74$, $df = 1,39$, $p = .194$, $R^2 < 0.003$) or interaction effect of

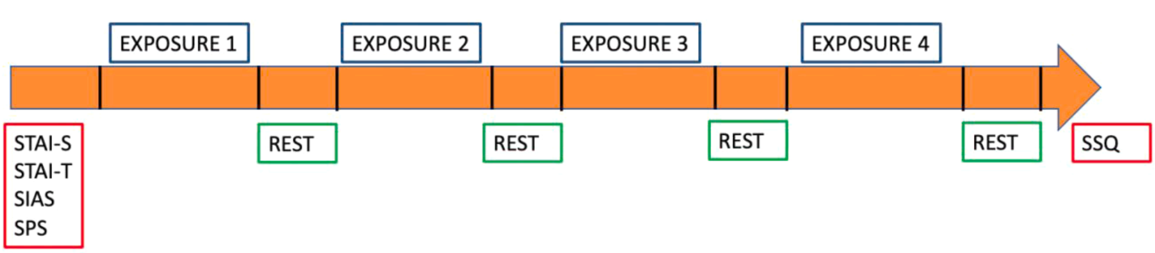

Fig. 4. Experimental timeline.

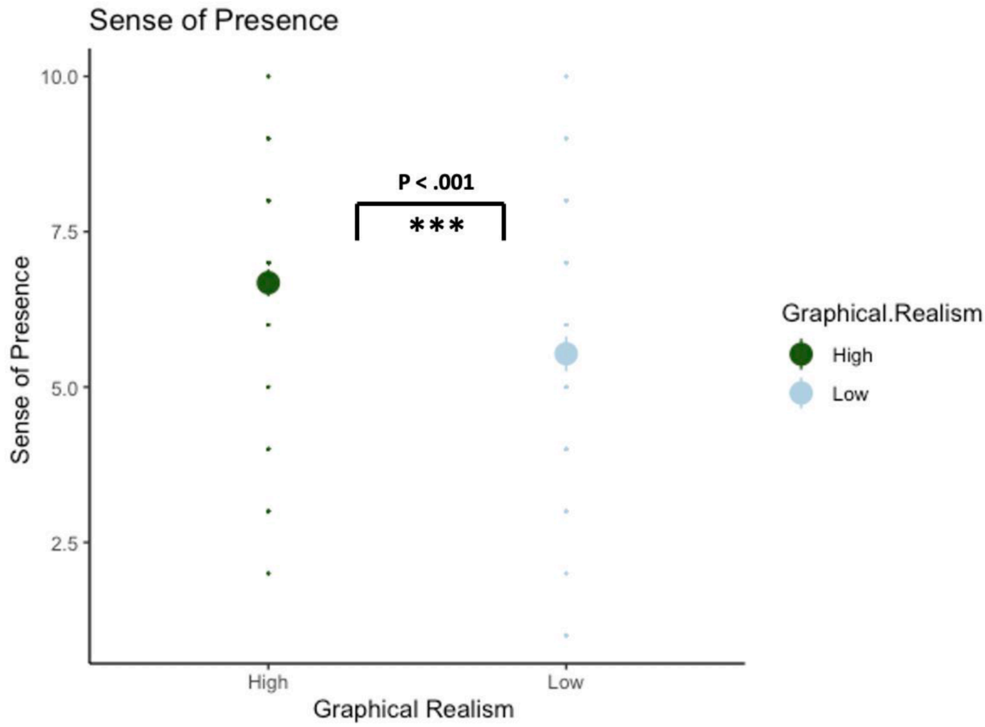

Fig. 5. Sense of presence in high vs. low graphical (x-axis) realism.

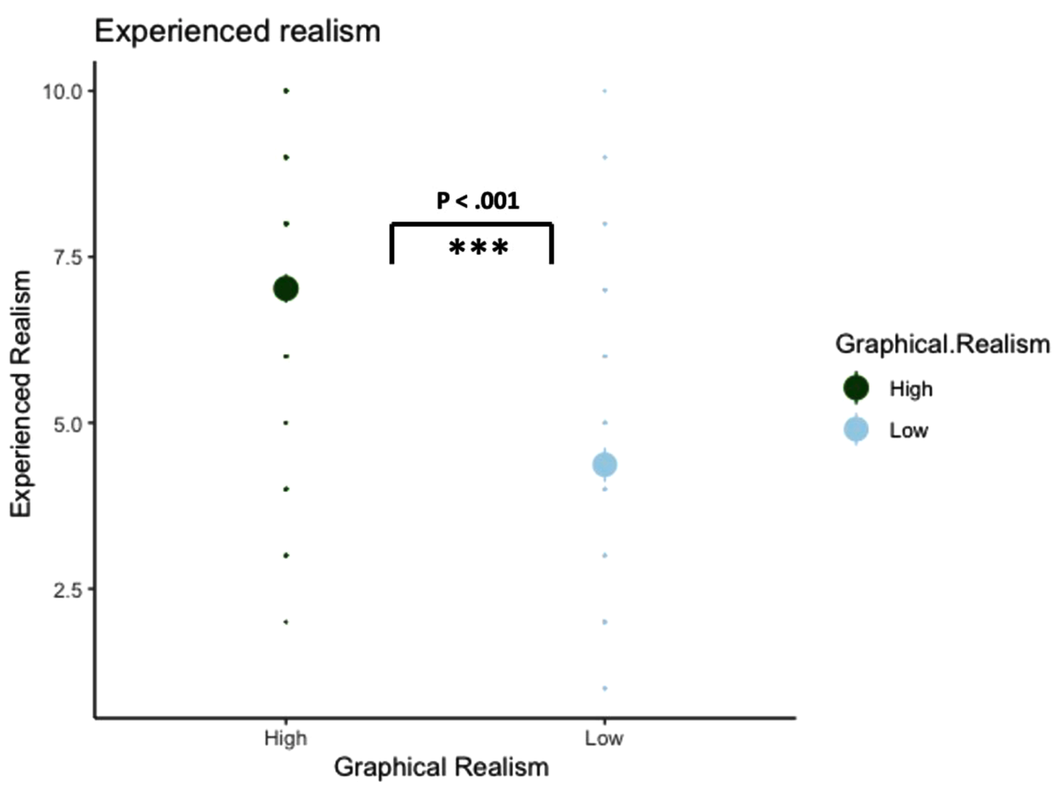

Fig. 6. Experienced realism in high vs. low graphical (x-axis) realism.

*graphical realism * acoustic realism* (*F* = *2.67, df* = *1,39, p* = *.109,* R^2 *< 0.003*) regarding the NS-SCR during public speaking were found. In order to understand the possible habituation effect due to repeated exposure, the exploratory analysis revealed a main effect of the *first exposure (F* = *21.41, df* = *1,39, p < .001,* R^2 *=0.039)* on NS-SCR, suggesting a higher number of non-specific skin conductance response during the first public speaking (*M* = 33.3, SD = 16.5) compared to the following performances (*M* = 26.6, SD = 16.0) (Fig. 9).

Skin conductance response (SCR): No main effect of *graphical realism (F* = *2.97, df* = *1,39 p* = *.092, ω2* = *0.016)* was found. A main effect of

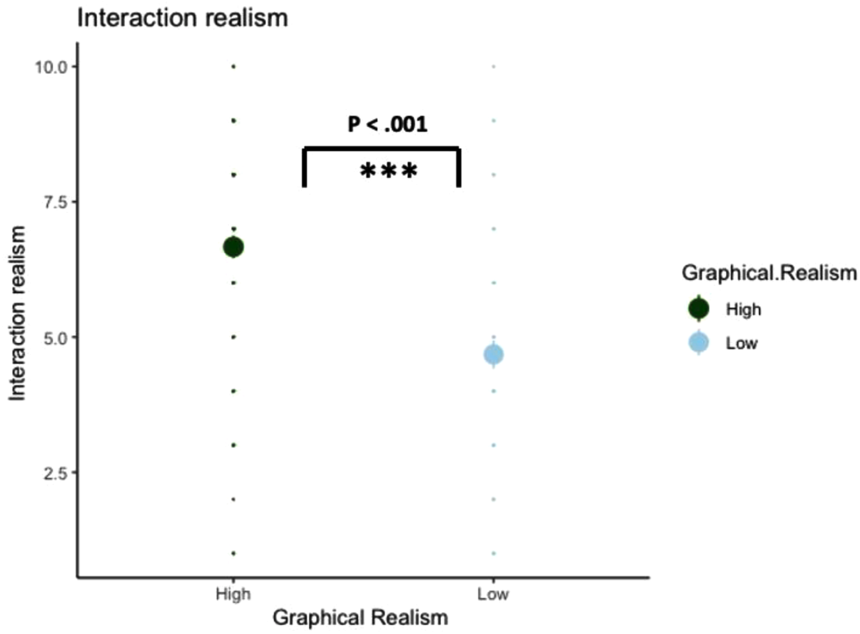

Fig. 7. Interaction realism in high vs. low graphical (x-axis)) realism.

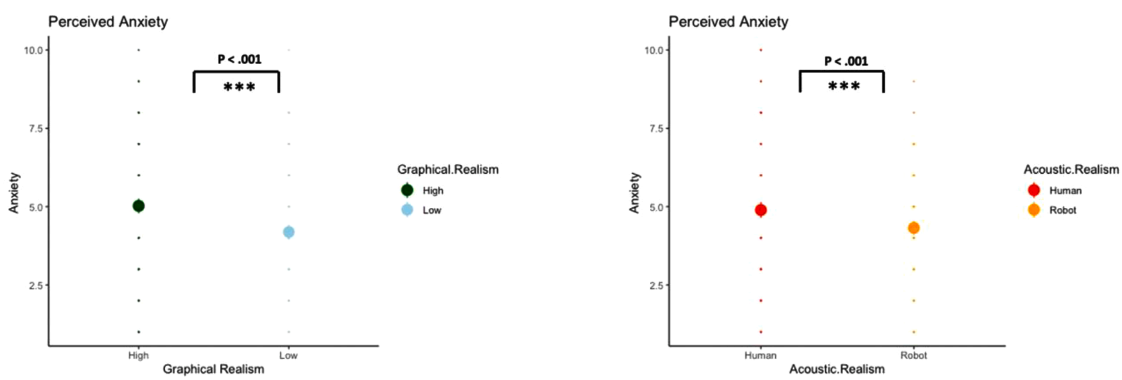

Fig. 8. Perceived anxiety in high vs. low graphical (left) and acoustic (right) realism.

acoustic realism ($F = 4.58$, $df = 1$, 39 $p = .045$, $\omega 2 = 0.023$) was found, indicating that the robotic voice evoked higher skin conductance response during the question ($M = 0.313$, SD $= 0.520$) compared to the human voice ($M = 0.192$, SD $= 0.229$). No interaction effect graphical realism * acoustic realism was found *($F = 1.93$, $df = 1$, 39 $p = .171$, $\omega 2 = 0.005$)* (Fig. 10).

3.4. Correlation matrix

3.4.1. Perceived anxiety subjective reports and physiological activity

A significant correlation was observed between perceived anxiety and reports of the virtual reality (VR) experience. Specifically, the sense of presence and experienced realism were positively correlated (r = 0.78, $p < .001$) (Fig. 11, *Panel A*). Additionally, both sense of presence (r = 0.38, $p < .001$) and experienced realism ($r = 0.40$, $p < .001$) were positively correlated with perceived anxiety (Fig. 11, *Panel B and C*). Regarding physiological activity, the number of NS-SCR increased with the perceived anxiety ($r = 0.35$, $p < .001$) experienced during the public speaking performance (Fig. 11, *Panel D*). A positive relationship between NS-SCR and the sense of presence ($r = 0.25$, $p = .002$) was also found.

4. Discussion

The main aim of this study was to understand the effects of graphic and acoustic features of the designed virtual audience on anxiety levels in a public speaking task. We measured participants' subjective reports concerning the perceived anxiety, the sense of presence, the sense of overall realism (experienced realism), and the realism of the interaction (interaction realism) in each simulated scenario (e.g., human vs. cartoon-like audience; human vs. robotic audience voice). Electrodermal activity was also recorded throughout all VR exposures (NS-SCR) and after the interactions with the audience (SCR) in order to analyze a more implicit measure of anxiety and presence beyond explicit self-assessment.

4.1. Anxiety levels in VR public speaking depend on the audience features

Concerning the participants' anxiety levels evoked by the task, the results showed significant effects of the graphic and auditory features of the audience. As expected, more realistic virtual reality scenarios, in terms of visual and auditory components, induced more anxiety compared to the less realistic audience. Specifically, the perceived anxiety increased in the case of the human audience (compared to the cartoon) and the use of human voice (compared to the robotic voice).

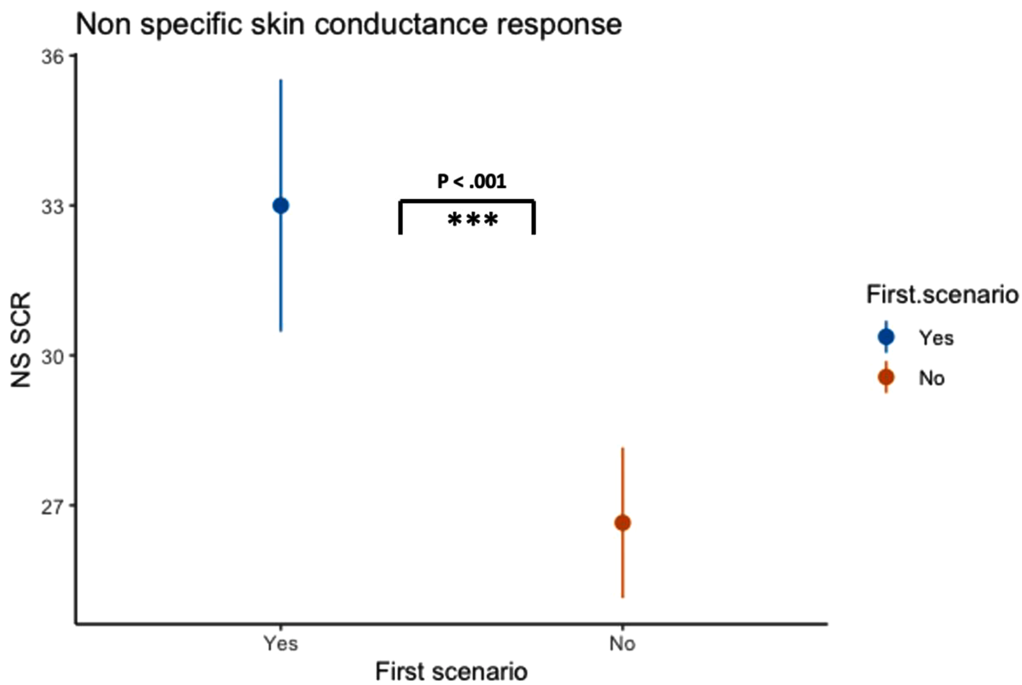

Fig. 9. Nonspecific skin conductance response across baseline and public speaking scenarios.

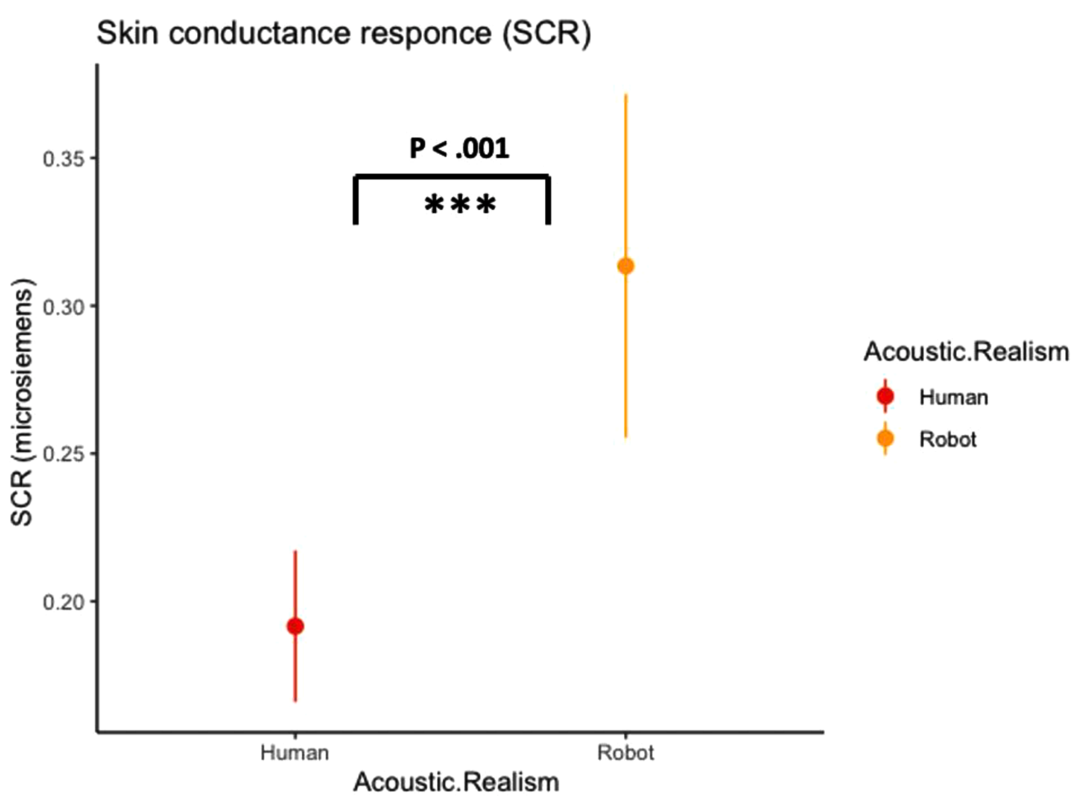

Fig. 10. SCR in high vs. low acoustic realism.

That is, both visual and acoustic aspects of the virtual audience quality affected self-reported anxiety during the exposure (see also (Kwon et al., 2013). The lack of interaction effect would seem to mean that the level of acoustic realism does not influence the effect of graphical realism on anxiety levels, and vice versa. Interestingly, the effect of audience design (both graphical and acoustic) did not result in significant changes in physiological arousal (NS-SCR) during the exposure. In this respect, a 'surprise effect' between the first scenario and those following may have affected the results. That is, most of the participants were not familiar with VR environments and the presence of a novel situation may have increased their overall level of arousal. In line with this assumption, the results showed a strong arousal response during the initial part of the overall experimental experience (i.e., a main effect of the first scenario in NS-SCR), while no differences due to graphical and acoustic were

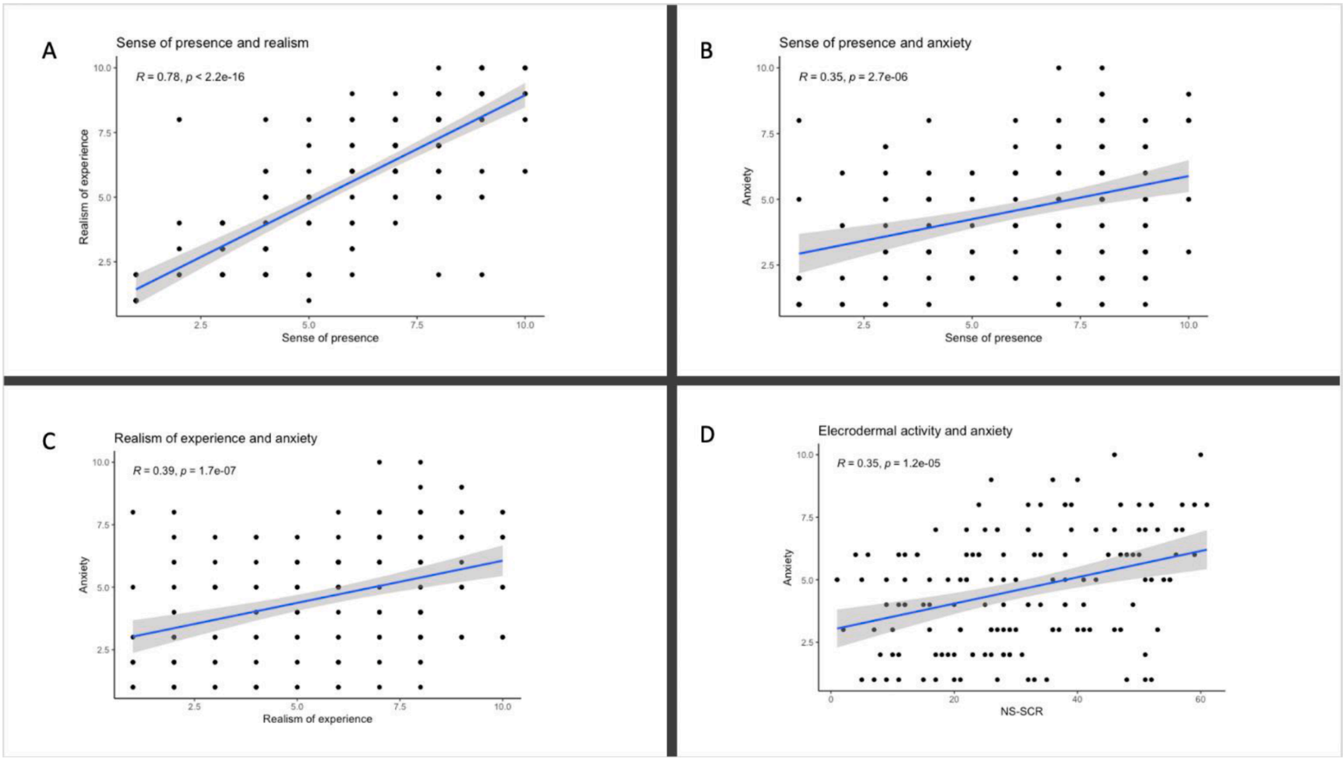

Fig. 11. Correlation matrix between perceived anxiety, subjective report, and NS-SCR.

found. Similarly to previous studies in the literature, repeated sessions of public speaking exposure in VR decrease anxiety level, probably due to the habituation effect against the stressful situation (Reeves et al., 2021; Takac et al., 2019). In order to avoid these possible confounding effects, future studies might use a between-group rather than repeated measures experimental design to explore the effect of graphical and acoustic realism on the physiological activation of public speaking performed in virtual reality.

This experiment was based on evidence suggesting that public speaking in VR offers a good opportunity to overcome the anticipation of stress and anxiety in human performance (Palmas et al., 2021; Poeschl, 2017; Premkumar et al., 2021; Stupar-Rutenfrans et al., 2017; Takac et al., 2019). As Klinger suggested, VR allows social anxiety to be induced like in a real-life social situation but in a more controlled environment (Klinger et al., 2005). The benefits of VR exposure training are strictly connected to facing fearful situations while avoiding adverse outcomes (Colombo et al., 2014; Nazir et al., 2012; Rothbaum et al., 2000). Apart from this, VR has an additional potential: the way in which the client/user is exposed to these fearful situations may easily be modulated, and the training can be adapted and customized based on the client's needs and individual differences. In the case of public speaking, and on the basis of our findings, the audience should be less realistic (i.e., less similar to their real counterpart) to avoid an excessively stressful situation. (Dafit Pitoyo and Asib, 2019; Pramana et al., 2018; Ryu et al., 2018; Sailer and Homner, 2020; Su, 2016).

4.2. Relationship between anxiety, realism and sense of presence

The experimental design was also conceptualized to investigate how each sensory component contributes to recreating a "realistic" virtual environment, in terms of self-report and physiological responses, in specific cases of stressful situations, such as public speaking. Understanding which perceptual elements are fundamental to creating a sense of realistic experience, and its relationship with the sense of presence, becomes a fundamental issue for the success of any training or clinical rehabilitation procedure in virtual reality. In fact, the sense of presence is a key concept for obtaining realistic responses in virtual environments, comparable to those occurring during real-world reactions (Felnhofer et al., 2014; Higuera-Trujillo et al., 2017; Ling et al., 2014; Slater, 2002; Slater et al., 2009a). However, the sense of presence does not necessarily coincide with experienced realism, or at least not without having a discussion on the definition of realism itself. One of the main issues that arises when people try to reproduce 'reality' is, in fact, the question about 'what is real'. What are the standards that humans use to define 'real'? Is it the exact physical correspondence between the simulated environment and the functioning of the real world (with the implementation of all physical laws discovered so far; something that at the moment is computationally and practically impossible), or is it about the perceptual level of similarity between the memory stored sensorial experiences of the word (mental schemes) and the perception of the simulated environment? And again, is realism related to the quality of the interaction rather than the sensory aspects of the environment? For this reason, it is important to investigate both presence and experienced realism in public speaking scenarios. In our hypothesis, experienced realism and sense of presence are connected to each other but are also related to the sensory aspects of the simulated VR environment.

Based on this study's results, graphical realism (being in front of a human avatar audience) significantly affected the sense of presence and the experienced realism during the task, while the acoustic aspect does not seem to play a crucial role in these measures. However, concerning the interaction with the audience (i.e., interaction realism), the acoustic features of the environment become relevant, indicating less perceived interaction realism in the case of a robotic voice. Indeed, despite the presence of a human avatar, the robotic voice negatively affected the evaluation of interaction realism with the audience. This effect is perhaps reflected also in the perceived anxiety reported by the participants, for which the human voice (high acoustic realism) induced higher anxiety compared to the robotic ones. Interestingly, after the presentation of the audience question, the robotic voice increased the SCR amplitude compared to the same question presented with the human voice. This result might be interpreted as a break in the sense of presence during the virtual experience, as previously reported in other studies (Slater et al., 2006a). In general, increasing the realism of the audience would directly affect the perceived anxiety during public speaking, probably through a modulation in the sense of presence. As Ling reported, the perceived anxiety and virtual reality quality experience reports seem to be reciprocally connected (Ling et al., 2014). As far as this point is concerned, our study showed a relationship between the sense of presence and experienced realism: the higher the experienced realism, the higher the sense of presence. Both realism and presence resulted in being positively correlated with perceived anxiety. In fact, a positive correlation between the sense of presence and non-specific skin conductance responses during the exposure was found, which might confirm the use of physiological arousal as a marker for presence in virtual environments (Gertler et al., 2020; Grassini and Laumann, 2020; Niles et al., 2015; Riva et al., 2003). As well, non-specific skin conductance responses resulted in being positively correlated with the anxiety reported by the participants.

The considerations emerging from this study might be taken as valid suggestions for training applications in virtual reality. For example, using a cartoon audience instead of real characters might be a powerful tool for learning social and individual skills, and more in general, for behavioral change to occur (Cleto, 2020; Pereira et al., 2014; Sailer and Homner, 2020). That is, creating a game where people have to interact, also verbally with cartoons might contribute to address public speaking issues through a less stressful approach in the overall population (Grivokostopoulou et al., 2017; Ryu et al., 2018). It seems clear that cartoon avatars appear as a lower-stress environment compared to a human avatar-based audience. Regarding this, Bailenson advanced the concept of transforming social interactions by means of computed mediated communication, to promote behavioral changes through environmental changes (Bailenson and Yee, 2008). In one of these studies (2005) eye gazing of virtual avatars was modified to increase a user's persuasiveness during a virtual reality listening task (Bailenson et al., 2005a). In the field of public speaking training, Bellido-Rivas and colleagues tested the effect of two-session speaking training in front of a cartoon audience while the participant embodied a cartoon avatar, compared to a group that embodied a human avatar and spoke in front of a human audience (Bellido Rivas et al., 2021). Both groups performed the last exposure in front of a human audience. The authors reported no effect of cartoon training, while the group exposed to the human avatar decreased their anxiety. However, this lack of improvement in anxiety state after the cartoon audience speech training in their study might be related to the fewer sessions of training adopted. In fact, in their protocol, the group exposed to the cartoon avatar shifted after just two training sessions into the human avatar condition, while the control group performed all three speaking conditions in front of a human avatar audience, whereby habituation effects were likely stronger. Moreover, the perceived anxiety was measured by state anxiety questionnaires only. It might be possible that the use of gamification strategies based on less realistic avatars (i.e., cartoons) for the modulation of public speaking anxiety may require more than two sessions in order to be effective, and/or the addition of further manipulation of avatar behavior (Pertaub et al., 2002; Premkumar et al., 2021). Undoubtedly, more evidence is required to fully understand the impact of less realistic avatars on public speaking anxiety.

It is important to considered that, in a multi users' simulation within virtual environments (human to human or human to AI avatar), another fundamental aspect that should be considered beside presence is the concept of 'social presence' (Blascovich et al., 2002; Bulu, 2012; Oh et al., 2018b; Yoon et al., 2019). Social presence may be defined as the "sense of being with another", for which user-avatar interactions in virtual environments "are felt to be real". Although in the present experiment, social presence was not directly investigated by means of specific questionnaires, we measured the interaction realism perceived after each exposure (the realism of the interaction with other avatars in the simulated scenario). We found that both acoustic and graphical realism increased the interaction realism perceived during the simulation, perhaps suggesting that social presence can be, at least in part, determined by the degree of (multisensory) realism of the virtual scenario, rather than by visual aspects only. However, given that only one statement on the interaction realism was used in this study and that a more direct assessment of social presence was not adopted (also due to the lack of a validated scale in the Italian language) we cannot draw any firm conclusions on this aspect. Moreover, in the present study, we used a robotic voice coupled with human vs. cartoon avatars to manipulate the acoustic realism of the audience. It is worth noting that the robotic voice used might be perceived as unfamiliar and impersonal to the participants. Conversely, a cartoon voice (e.g., coupled with the cartoon avatar with a voice more similar to the one used in film animation) might affect the perceived anxiety and realism differently. In a previous study, Higgins, et al., investigated how synthetic voice affects social presence, emotional reaction and discomfort toward virtual avatar, in immersive and non-immersive environments (Higgins et al., 2022). The authors reported no effect on social presence and a partial modulation of emotional reaction using a synthetic (human-based) voice in the virtual environment condition. On the contrary, social presence decreased if the interaction occurred in a non-immersive environment where a synthetic voice based on the text-to-speech system was presented (i.e., lower acoustic realism level than the human-based robotic voice) (Zibrek et al., 2021). That is, sound design is becoming an exciting aspect to understand better the relationship between experience, presence, and anxiety within public speaking tasks performed in a virtual environment (Gabory and Chollet, 2020). Further research should then be addressed to the question of how our perception weights acoustic and visual information in virtual environments and to define the amount of mismatch between visual and acoustic information that can be tolerated by the participants in order to perceive the virtual environment as realistic (also from a social point of view). We believe that the topic related to the role of multisensory cues in determining social presence (and its relationship with perceived anxiety in virtual public speaking) should be carefully considered in future studies. In fact, the possibility of interacting in social worlds through virtual reality – metaverses – may offer a very promising field of study in the years to come. As the last point, the present study did not directly manipulate the behavioral realism of the audience, keeping uninvestigated if and how behavioral realism interacts with the graphical and acoustic features in VR public speaking simulation (Guadagno et al., 2007; Herrera et al., 2020). In our simulation, both human and cartoon avatars express movement and gesture (i.e., head movement, hand movement, eye blink, eye gazing, applying as similar as possible to the cartoon audience) during the participant's speech to increase the fidelity of the virtual simulation, also to avoid confounding effect driven by different level of behavioral realism. It is safe to assume that the high level of behavioral realism might have increased the perceived anxiety, the realism experienced, and the presence, presumably in the same way for all the simulations presented to the participants. An interesting question is whether and how graphical and acoustic realism impact the perceived anxiety without the presence of realistic behavior by the audience since it has been suggested a sort of "congruency effect" between graphical and behavioral realism (i.e., lack of gestures or incongruent behavior might impact more in case of a human avatar than cartoon avatar, for which no specific behavior is expected) (Bailenson et al., 2005b; Chollet and Scherer, 2017), future investigations need to address how graphical and acoustic realism effect depends by the behavioral realism expressed during the virtual social interaction.

5. Limits of the study

The findings of this study should be considered in light of certain limitations. Firstly, the study included only healthy (young) participants, and we cannot exclude the possibility that clinical participants might be differently affected by the realism of the scene. That is, patients might be more sensitive than other people to even lower levels of realism. This important topic should also be addressed in the years to come. Additionally, the speech argument used in our study (a recipe) was not so difficult or personally relevant (e.g., compared to talking about more private or personal topics; or topics that are considered more anxiogenics by the participant). It could be interesting, in future studies, to see how the graphical and acoustic aspect affects perceived anxiety in the case of more personally-relevant speeches. Moreover, the study did not find any difference in terms of physiological responses with respect to the audience features. This is probably related to a "ceiling effect" due to the strong arousal activation elicited by the first experience, which masked any possible difference in the scenario presented. Future studies should address this question by means of a between-subjects design across conditions.

Another limit is related to the lack of more standardized measures of presence (and social presence) after each exposure. We did not use self-assessment questionnaires for these variables after each condition in order to avoid experiment overlength and participants' fatigue. Given that our primary outcome was the perceived anxiety during the public speaking task, we avoided using a long questionnaire session that could have affected the performance of the participants. The lack of a standardized version of the questionnaire in the Italian language compromised the possibility of using specific subscales for the purpose of analyzing further presence and social presence. Given the importance of the social presence concept inside the interactive virtual environment, future studies will need to accurately explore the relationship between realism and social presence within public speaking scenarios.

6. Conclusions

In conclusion, this study showed that the audience features, in terms of graphical and acoustic realism, can modulate the perceived anxiety during a virtual reality public speaking exposure. Increasing the audience's multisensory realism directly affects the capabilities of a virtual environment to induce public speaking anxiety. This modulation would seem to be related to the sense of presence, the experienced realism, and the realism of the interaction experienced by participants within the virtual environment. We also showed that participants' anxiety levels correlated with the sense of presence and the perception of realism in the simulated environment. Moreover, we found a positive correlation between physiological arousal (NS-SCR) and anxiety levels, sense of presence (the more presence, the more anxiety, the more physiological arousal). Taken as a whole, the results of the present experiment may be taken to suggest the importance of investigating both the visual and acoustic properties of an environment in order to modulate participants' anxiety in a possible exposure therapy or training application.

CRediT authorship contribution statement

Matteo Girondini: Conceptualization, Methodology, Data curation, Writing – original draft, Writing – review & editing. **Milena Stefanova:** Software, Methodology. **Margherita Pillan:** Conceptualization, Writing – review & editing. **Alberto Gallace:** Conceptualization, Methodology, Writing – original draft, Writing – review & editing, Supervision.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability

Data will be made available on request.

References

Aderka, I.M., Hofmann, S.G., Nickerson, A., Hermesh, H., Gilboa-Schechtman, E., Marom, S., 2012. Functional impairment in social anxiety disorder. J. Anxiety Disord. 26, 393–400. https://doi.org/10.1016/J.JANXDIS.2012.01.003.

Åhs, F., Sollers, J.J., Furmark, T., Fredrikson, M., Thayer, J.F., 2009. High-frequency heart rate variability and cortico-striatal activity in men and women with social phobia. Neuroimage 47, 815–820. https://doi.org/10.1016/J.NEUROIMAGE.2009.05.091.

Allmendinger, K., 2010. Social presence in synchronous virtual learning situations: the role of nonverbal signals displayed by avatars. Educ. Psychol. Rev. https://doi.org/10.1007/s10648-010-9117-8.

Arsalan, A., Majid, M., 2021. Human stress classification during public speaking using physiological signals. Comput. Biol. Med. 133, 104377 https://doi.org/10.1016/J.COMPBIOMED.2021.104377.

Bailenson, J.N., Beall, A.C., Loomis, J., Blascovich, J., Turk, M., 2005a. Transformed social interaction, augmented gaze, and social influence in immersive virtual environments. Hum. Commun. Res. 31, 511–537. https://doi.org/10.1111/j.1468-2958.2005.tb00881.x.

Bailenson, J.N., Swinth, K., Hoyt, C., Persky, S., Dimov, A., Blascovich, J., 2005b. The independent and interactive effects of embodied-agent appearance and behavior on self-report, cognitive, and behavioral markers of copresence in immersive virtual environments. Presence Teleoperators Virtual Environ. 14, 379–393. https://doi.org/10.1162/105474605774785235.

Bailenson, J.N., Yee, N., 2008. Virtual interpersonal touch: haptic interaction and copresence in collaborative virtual environments. Multimed. Tools Appl. 37, 5–14. https://doi.org/10.1007/s11042-007-0171-2.

Bassett, J.R., Marshall, P.M., Spillane, R., 1987. The physiological measurement of acute stress (public speaking) in bank employees. Int. J. Psychophysiol. 5, 265–273. https://doi.org/10.1016/0167-8760(87)90058-4.

Beckler, K., 2010. State-Trait Anxiety Inventory For Adults Sampler Set Manual, Instrument and Scoring Guide. 1983 Consult. Psychol. Press. Inc. Mind Gard. Inc, pp. 0–78. https://doi.org/10.1037/T06496-000.

Behnke, R.R., Sawyer, C.R., 2001. Public speaking arousal as a function of anticipatory activation and autonomic reactivity. Int. J. Phytoremediation 21, 73–85. https://doi.org/10.1080/08934210109367740.

Bellido Rivas, A.I., Navarro, X., Banakou, D., Oliva, R., Orvalho, V., Slater, M., 2021. The influence of embodiment as a cartoon character on public speaking anxiety. Front. Virtual Real. 2 https://doi.org/10.3389/FRVIR.2021.695673/PDF.

Biocca, F., Harms, C., Burgoon, J.K., 2003. Toward a more robust theory and measure of social presence: review and suggested criteria. Presence Teleoperators Virtual Environ 12, 456–480. https://doi.org/10.1162/105474603322761270.

Birkett, M.A., 2011. The Trier Social Stress Test protocol for inducing psychological stress. J. Vis. Exp. 1–6. https://doi.org/10.3791/3238.

Blascovich, J., Loomis, J., Beall, A.C., Swinth, K.R., Hoyt, C.L., Bailenson, J.N., 2002. Immersive virtual environment technology as a methodological tool for social psychology. Psychol. Inq. 13, 103–124. https://doi.org/10.1207/S15327965PLI1302_01.

Bouchard, S., Dumoulin, S., Robillard, G., Guitard, T., Klinger, E., Forget, H., Loranger, C., Roucaut, F.X., 2017. Virtual reality compared with in vivo exposure in the treatment of social anxiety disorder: a three-arm randomised controlled trial. Br. J. Psychiatry 210, 276–283. https://doi.org/10.1192/bjp.bp.116.184234.

Bulu, S.T., 2012. Place presence, social presence, co-presence, and satisfaction in virtual worlds. Comput. Educ. 58, 154–161. https://doi.org/10.1016/j.compedu.2011.08.024.

Bush, J., 2008. Viability of virtual reality exposure therapy as a treatment alternative. Comput. Human Behav. 24, 1032–1040. https://doi.org/10.1016/j.chb.2007.03.006.

Bystrom, K.E., Barfield, W., Hendrix, C., 1999. A conceptual model of the sense of presence in virtual environments. Presence Teleoperators Virtual Environ. https://doi.org/10.1162/105474699566107.

Carl, E., Stein, A.T., Levihn-Coon, A., Pogue, J.R., Rothbaum, B., Emmelkamp, P., Asmundson, G.J.G., Carlbring, P., Powers, M.B., 2019. Virtual reality exposure therapy for anxiety and related disorders: a meta-analysis of randomized controlled trials. J. Anxiety Disord. 61, 27–36. https://doi.org/10.1016/j.janxdis.2018.08.003.

Carrillo, E., Moya-Albiol, L., González-Bono, E., Salvador, A., Ricarte, J., Gómez-Amor, J., 2001a. Gender differences in cardiovascular and electrodermal responses to public speaking task: the role of anxiety and mood states. Int. J. Psychophysiol. 42, 253–264. https://doi.org/10.1016/S0167-8760(01)00147-7.

Carrillo, E., Moya-Albiol, L., González-Bono, E., Salvador, A., Ricarte, J., Gómez-Amor, J., 2001b. Gender differences in cardiovascular and electrodermal responses to public speaking task: the role of anxiety and mood states. Int. J. Psychophysiol. 42, 253–264. https://doi.org/10.1016/S0167-8760(01)00147-7.

Chaby, L.E., Sheriff, M.J., Hirrlinger, A.M., Braithwaite, V.A., 2015. Can we understand how developmental stress enhances performance under future threat with the Yerkes-Dodson law? Commun. Integr. Biol. 8 https://doi.org/10.1080/19420889.2015.1029689.

Chollet, M., Neubauer, C., Ghate, P., Scherer, S., 2018. Influence of individual differences when training public speaking with virtual audiences. In: Proceedings of the 18th International Conference on Intelligent Virtual Agents, IVA 2018. New York, NY, USA. ACM, pp. 1–7. https://doi.org/10.1145/3267851.3267874.

Chollet, M., Scherer, S., 2017. Perception of virtual audiences. IEEE Comput. Graph. Appl. 37, 50–59. https://doi.org/10.1109/MCG.2017.3271465.

Cleto, B., 2020. Learning Systems and Gamification. IGI Global, pp. 54–67. https://doi.org/10.4018/978-1-7998-7472-0.ch004. https://Services.Igi-Global.Com/Resolvedoi/Resolve.Aspx?Doi=10.4018/978-1-7998-7472-0.Ch004.

Colombo, S., Nazir, S., Gallace, A., Manca, D., 2014. Experiment-based decision making in complex systems. Chem. Eng. Trans. 36, 85–90. https://doi.org/10.3303/CET1436015.

Connolly, J.F., Frith, C.D., 1978. Effects of stimulus variability on the amplitude and habituation of the electrodermal orienting response. Psychophysiology 15, 550–555. https://doi.org/10.1111/j.1469-8986.1978.tb03108.x.

Croft, R.J., Gonsalvez, C.J., Gander, J., Lechem, L., Barry, R.J., 2004. Differential relations between heart rate and skin conductance, and public speaking anxiety. J. Behav. Ther. Exp. Psychiatry 35, 259–271. https://doi.org/10.1016/J.JBTEP.2004.04.012.

Dafit Pitoyo, M., Asib, A., 2019. Gamification based assessment: a test anxiety reduction through game elements in quizizz platform. IJER 4, 22–32.

de Melo, C.M., Carnevale, P.J., Gratch, J., 2013. Agent or avatar? Using virtual confederates in conflict management research. Acad. Manag. Proc. 2013, 10859. https://doi.org/10.5465/ambpp.2013.10859abstract.

Dimberg, U., Fredrikson, M., Lundquist, O., 1986. Autonomic reactions to social and neutral stimuli in subjects high and low in public speaking fear. Biol. Psychol. 23, 223–233. https://doi.org/10.1016/0301-0511(86)90001-3.

Dionisio, J.D.N., Burns, W.G., Gilbert, R., 2013. 3D virtual worlds and the metaverse: current status and future possibilities. ACM Comput. Surv. 45 https://doi.org/10.1145/2480741.2480751.

Dubosc, C., Gorisse, G., Christmann, O., Fleury, S., Poinsot, K., Richir, S., 2021. Impact of avatar facial anthropomorphism on body ownership, attractiveness and social presence in collaborative tasks in immersive virtual environments. Comput. Graph. 101, 82–92. https://doi.org/10.1016/j.cag.2021.08.011.

Ebrahimi, O.V., Pallesen, S., Kenter, R.M.F., Nordgreen, T., 2019. Psychological interventions for the fear of public speaking: a meta-analysis. Front. Psychol. 10, 488. https://doi.org/10.3389/FPSYG.2019.00488/BIBTEX.

El-Yamri, M., Romero-Hernandez, A., Gonzalez-Riojo, M., Manero, B., 2019. Designing a VR game for public speaking based on speakers features: a case study. Smart Learn. Environ. 6, 1–15. https://doi.org/10.1186/S40561-019-0094-1, 2019 61.

Faul, F., Erdfelder, E., Lang, A.G., Buchner, A., 2007. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav. Res. Methods 392 (39), 175–191. https://doi.org/10.3758/BF03193146, 2007.

Felnhofer, A., Kothgassner, O.D., Hetterle, T., Beutl, L., Hlavacs, H., 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. Cyberpsychology, Behav. Soc. Netw. 17, 310–316. https://doi.org/10.1089/CYBER.2013.0472/ASSET/IMAGES/LARGE/FIGURE3.JPEG.

Ferreira Marinho, A.C., Mesquita de Medeiros, A., Côrtes Gama, A.C., Caldas Teixeira, L., 2017. Fear of public speaking: perception of college students and correlates. J. Voice 31, 127.e7–127.e11. https://doi.org/10.1016/J.JVOICE.2015.12.012.

Fich, L.B., Jönsson, P., Kirkegaard, P.H., Wallergård, M., Garde, A.H., Hansen, Å., 2014. Can architectural design alter the physiological reaction to psychosocial stress? A virtual TSST experiment. Physiol. Behav. 135, 91–97. https://doi.org/10.1016/J.PHYSBEH.2014.05.034.

Finn, A.N., Sawyer, C.R., Behnke, R.R., 2009. A model of anxious arousal for public speaking. Commun. Educ. 58, 417–432. https://doi.org/10.1080/03634520802268891.

Frassinetti, F., Bolognini, N., Làdavas, E., 2002. Enhancement of visual perception by crossmodal visuo-auditory interaction. Exp. Brain Res. 147, 332–343. https://doi.org/10.1007/s00221-002-1262-y.

Gabory, E., Chollet, M., 2020. Investigating the influence of sound design for inducing anxiety in virtual public speaking. In: ICMI 2020 Companion - Companion Publication of the 2020 International Conference on Multimodal Interaction. ACM, pp. 492–496. https://doi.org/10.1145/3395035.3425227.

Gallace, A., 2022. Haptic interaction in virtual reality: are we ready for the metaverse? Neuroscientific and behavioral considerations. Handbook of Research on Implementing Digital Reality and Interactive Technologies to Achieve Society 5.0. IGI Global, pp. 1–14. https://doi.org/10.4018/978-1-6684-4854-0.CH001.

Gallace, A., Ngo, M.K., Sulaitis, J., Spence, C., 2011. Multisensory presence in virtual reality: possibilities & limitations. Multiple Sensorial Media Advances and Applications: New Developments in MulSeMedia, pp. 1–38. https://doi.org/10.4018/978-1-60960-821-7.ch001.

Garau, M., 2003. The impact of avatar fidelity on social interaction in virtual environments.

Geller, T., 2008. Overcoming the uncanny valley. IEEE Comput. Graph. Appl. 28, 11–17. https://doi.org/10.1109/MCG.2008.79.

Gertler, J., Novotny, S., Poppe, A., Chung, Y.S., Gross, J.J., Pearlson, G., Stevens, M.C., 2020. Neural correlates of non-specific skin conductance responses during resting state fMRI. Neuroimage 214, 116721. https://doi.org/10.1016/j.neuroimage.2020.116721.

Giesen, J.M., McGlynn, F.D., 1977. Skin conductance and heart-rate responsivity to public speaking imagery among students with high and low self-reported fear: a comparative analysis of "response" definitions. J. Clin. Psychol. 33, 68–76. https://doi.org/10.1002/1097-4679(197701)33:1+<68::AID-JCLP2270330114>3.0.CO;2-P.

Girondini, M., Stefanova, M., Pillan, M., Gallace, A., 2023. The effect of previous exposure on virtual reality induced public speaking anxiety: a physiological and behavioral study. Cyberpsychology, Behav. Soc. Netw. 26, 127–133. https://doi.org/10.1089/cyber.2022.0121.

Goncalves, G., Monteiro, P., Coelho, H., Melo, M., Bessa, M., 2021. Systematic review on realism research methodologies on immersive virtual, augmented and mixed realities. IEEE Access. https://doi.org/10.1109/ACCESS.2021.3089946.

Gonzalez-Bono, E., Moya-Albiol, L., Salvador, A., Carrillo, E., Ricarte, J., Gomez-Amor, J., 2002. Anticipatory autonomic response to a public speaking task in women: the role of trait anxiety. Biol. Psychol. 60, 37–49. https://doi.org/10.1016/S0301-0511(02)00008-X.

Goodman, W.K., Janson, J., Wolf, J.M., 2017. Meta-analytical assessment of the effects of protocol variations on cortisol responses to the Trier Social Stress Test. Psychoneuroendocrinology 80, 26–35. https://doi.org/10.1016/J.PSYNEUEN.2017.02.030.

Grassini, S., Laumann, K., 2020. Questionnaire measures and physiological correlates of presence: a systematic review. Front. Psychol. 0, 349. https://doi.org/10.3389/FPSYG.2020.00349.

Green, C.D., Yerkes, R.M., Dodson, J.D., 1908. The relation of strength of stimulus to rapidity of habit-formation. J. Comp. Neurol. Psychol. 18, 459–482.

Grivokostopoulou, F., Perikos, I., Hatzilygeroudis, I., 2017. An innovative educational environment based on virtual reality and gamification for learning search algorithms. In: Proceedings - IEEE 8th International Conference on Technology for Education, T4E 2016. Institute of Electrical and Electronics Engineers Inc., pp. 110–115. https://doi.org/10.1109/T4E.2016.029

Guadagno, R.E., Blascovich, J., Bailenson, J.N., Mccall, C., 2007. Virtual humans and persuasion: the effects of agency and behavioral realism. Media Psychol 10, 1–22, 10.108/15213260701300865.

Guimarães, M., Prada, R., Santos, P.A., Dias, J., Jhala, A., Mascarenhas, S., 2020. The impact of virtual reality in the social presence of a virtual agent. In: Proceedings of the 20th ACM International Conference on Intelligent Virtual Agents, IVA 2020. Association for Computing Machinery, Inc. https://doi.org/10.1145/3383652.3423879.

Harris, Sandra R, Kemmerling, R.L., North, M.M., 2002. Brief virtual reality therapy for public speaking anxiety. Cyberpsychology Behav. 5, 543–550. https://doi.org/10.1089/109493102321018187.

Harris, V.A., Katkin, E.S., 1975. Primary and secondary emotional behavior: an analysis of the role of autonomic feedback on affect, arousal, and attribution. Psychol. Bull. 82, 904–916. https://doi.org/10.1037/0033-2909.82.6.904.

Heimberg, R.G., 2002. Cognitive-behavioral therapy for social anxiety disorder: current status and future directions. Biol. Psychiatry 51, 101–108. https://doi.org/10.1016/S0006-3223(01)01183-0.

Heimberg, R.G., Mueller, G.P., Holt, C.S., Hope, D.A., Liebowitz, M.R., 1992. Assessment of anxiety in social interaction and being observed by others: the social interaction anxiety scale and the Social Phobia Scale. Behav. Ther. 23, 53–73. https://doi.org/10.1016/S0005-7894(05)80308-9.

Herrera, F., Oh, S.Y., Bailenson, J.N., 2020. Effect of behavioral realism on social interactions inside collaborative virtual environments. Presence Teleoperators Virtual Environ 27, 163–182. https://doi.org/10.1162/PRES_a_00324.

Higgins, D., Zibrek, K., Cabral, J., Egan, D., McDonnell, R., 2022. Sympathy for the digital: influence of synthetic voice on affinity, social presence and empathy for photorealistic virtual humans. Comput. Graph. 104, 116–128. https://doi.org/10.1016/j.cag.2022.03.009.

Higuera-Trujillo, J.L., López-Tarruella Maldonado, J., Llinares Millán, C., 2017. Psychological and physiological human responses to simulated and real environments: a comparison between photographs, 360° panoramas, and virtual reality. Appl. Ergon. 65, 398–409. https://doi.org/10.1016/J.APERGO.2017.05.006.

Hinojo-Lucena, F.J., Aznar-Díaz, I., Cáceres-Reche, M.P., Trujillo-Torres, J.M., Romero-Rodríguez, J.M., 2020. Virtual reality treatment for public speaking anxiety in students. Advancements and results in personalized medicine. J. Pers. Med. 10 https://doi.org/10.3390/jpm10010014.

Ho, C.C., MacDorman, K.F., 2017. Measuring the uncanny valley effect: refinements to indices for perceived humanness, attractiveness, and eeriness. Int. J. Soc. Robot. 9, 129–139. https://doi.org/10.1007/s12369-016-0380-9.

Hoffman, H.G., Richards, T.L., Coda, B., Bills, A.R., Blough, D., Richards, A.L., Sharar, S.R., 2004. Modulation of thermal pain-related brain activity with virtual reality: evidence from fMRI. Neuroreport 15, 1245–1248. https://doi.org/10.1097/01.wnr.0000127826.73576.91.

Insko, B.E., 2003. Being There: concepts, effects and measurement of user presence in synthetic environments G 7 Measuring Presence: subjective, Behavioral and Physiological Methods.

Kahlon, S., Lindner, P., 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. https://doi.org/10.1186/S13034-019-0307-Y/FIGURES/3.

Kampmann, I.L., Emmelkamp, P.M.G., Hartanto, D., Brinkman, W.P., Zijlstra, B.J.H., Morina, N., 2016. Exposure to virtual social interactions in the treatment of social anxiety disorder: a randomized controlled trial. Behav. Res. Ther. 77, 147–156. https://doi.org/10.1016/J.BRAT.2015.12.016.

Kennedy, R.S., Lane, N.E., Berbaum, K.S., Lilienthal, M.G., 2009. Simulator sickness questionnaire: an enhanced method for quantifying simulator sickness. https://doi.org/10.1207/s15327108ijap0303_3 3, 203–220. https://doi.org/10.1207/S15327108IJAP0303_3.

Klinger, E., Bouchard, S., Légeron, P., Roy, S., Lauer, F., Chemin, I., Nugues, P., 2005. Virtual reality therapy versus cognitive behavior therapy for social phobia: a preliminary controlled study. Cyberpsychol. Behav. 8, 76–88. https://doi.org/10.1089/cpb.2005.8.76.

Kothgassner, O.D., Felnhofer, A., Hlavacs, H., Beutl, L., Palme, R., Kryspin-Exner, I., Glenk, L.M., 2016. Salivary cortisol and cardiovascular reactivity to a public speaking task in a virtual and real-life environment. Comput. Human Behav. 62, 124–135. https://doi.org/10.1016/J.CHB.2016.03.081.

Kwon, J.H., Powell, J., Chalmers, A., 2013. How level of realism influences anxiety in virtual reality environments for a job interview. Int. J. Hum. Comput. Stud. 71, 978–987. https://doi.org/10.1016/j.ijhcs.2013.07.003.

Kyrlitsias, C., Michael-Grigoriou, D., 2022. Social interaction with agents and avatars in immersive virtual environments: a survey. Front. Virtual Real. https://doi.org/10.3389/frvir.2021.786665.

Lacy, C.R., Contracta, R.J., Robbins, M.L., Tannenbaum, A.K., Moreyra, A.E., Chelton, S., Kostis, J.B., 1995. Coronary vasoconstriction induced by mental stress (simulated public speaking). Am. J. Cardiol. 75, 503–505. https://doi.org/10.1016/S0002-9149(99)80590-6.

Lanier, J., 2017. Dawn of the new everything : encounters with reality and virtual reality.

Lawler, K.A., 1980. Cardiovascular and electrodermal response patterns in heart rate reactive individuals during psychological stress. Psychophysiology 17, 464–470. https://doi.org/10.1111/j.1469-8986.1980.tb00185.x.

Ling, Y., Nefs, H.T., Morina, N., Heynderickx, I., 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. https://doi.org/10.1371/journal.pone.0096144.

MacIntyre, V.A., MacIntyr, P.D., Carre, G., 2010. Heart rate variability as a predictor of speaking anxiety. Commun. Res. Reports 27, 286–297. https://doi.org/10.1080/08824096.2010.496323.

Maffei, L., Masullo, M., Pascale, A., Ruggiero, G., Romero, V.P., 2016. Immersive virtual reality in community planning: acoustic and visual congruence of simulated vs real world. Sustain. Cities Soc. 27, 338–345. https://doi.org/10.1016/j.scs.2016.06.022.

Marucci, M., Di Flumeri, G., Borghini, G., Sciaraffa, N., Scandola, M., Pavone, E.F., Babiloni, F., Betti, V., Aricò, P., 2021. The impact of multisensory integration and perceptual load in virtual reality settings on performance, workload and presence. Sci. Reports 111 (11), 1–15. https://doi.org/10.1038/s41598-021-84196-8, 2021.

McDonnell, R., Breidty, M., Bülthoff, H.H., 2012. Render me real? Investigating the effect of render style on the perception of animated virtual humans. ACM Trans. Graph. 31 https://doi.org/10.1145/2185520.2185587.

McKinney, M.E., Gatchel, R.J., 1982. The comparative effectiveness of heart rate biofeedback, speech skills training, and a combination of both in treating public-speaking anxiety. Biofeedback Self. Regul. 7, 71–87. https://doi.org/10.1007/BF00999056.

Mitchell, W.J., Szerszen, K.A., Lu, A.S., Schermerhorn, P.W., Scheutz, M., MacDorman, K. F., 2011. A mismatch in the human realism of face and voice produces an uncanny valley. Iperception. https://doi.org/10.1068/i0415.

Mori, M., MacDorman, K.F., Kageki, N., 2012. The uncanny valley. IEEE Robot. Autom. Mag. 19, 98–100. https://doi.org/10.1109/MRA.2012.2192811.

Nass, C., Steuer, J., Tauber, E.R., 1994. are Social Actors. Hum. Factors 72–78.

Nazir, S., Gallace, A., Bordegoni, M., Colombo, S., Manca, D., 2012. Performance comparison of different training methods for industrial operators, in: Human Factors of Systems and Technology. pp. 235–242.

Nikula, R., 1991. Psychological correlates of nonspecific skin conductance responses. Psychophysiology 28, 86–90. https://doi.org/10.1111/j.1469-8986.1991.tb03392.x.

Niles, A.N., Craske, M.G., Lieberman, M.D., Hur, C., 2015. Affect labeling enhances exposure effectiveness for public speaking anxiety. Behav. Res. Ther. 68, 27–36. https://doi.org/10.1016/j.brat.2015.03.004.

Norman, G., 2010. Likert scales, levels of measurement and the "laws" of statistics. Adv. Heal. Sci. Educ. 15, 625–632. https://doi.org/10.1007/s10459-010-9222-y.

North, M.M., North, S.M., Coble, J.R., 1998. Virtual reality therapy: an effective treatment for the fear of public speaking. Int. J. Virtual Real. 3, 1–6. https://doi.org/10.20870/IJVR.1998.3.3.2625.

Nowak, K.L., Biocca, F., 2003. The effect of the agency and anthropomorphism on users' sense of telepresence, copresence, and social presence in virtual environments. Presence: Teleoperators and Virtual Environments. MIT Press, pp. 481–494. https://doi.org/10.1162/105474603322761289.

Oh, C.S., Bailenson, J.N., Welch, G.F., 2018a. A systematic review of social presence: definition, antecedents, and implications. Front. Robot. AI 5, 114. https://doi.org/10.3389/FROBT.2018.00114/BIBTEX.

Oh, C.S., Bailenson, J.N., Welch, G.F., 2018b. A systematic review of social presence: definition, antecedents, and implications. Front. Robot. AI. https://doi.org/10.3389/frobt.2018.00114.

Owens, M.E., Beidel, D.C., 2015. Can virtual reality effectively elicit distress associated with social anxiety disorder? J. Psychopathol. Behav. Assess. 37, 296–305. https://doi.org/10.1007/S10862-014-9454-X/TABLES/3.

Palmas, F., Reinelt, R., Cichor, J.E., Plecher, D.A., Klinker, G., 2021. Virtual reality public speaking training: experimental evaluation of direct feedback technology acceptance. In: Proceedings - 2021 IEEE Conference on Virtual Reality and 3D User Interfaces, VR 2021, pp. 463–472. https://doi.org/10.1109/VR50410.2021.00070.

Pan, X., Banakou, D., Slater, M., 2011. Computer based video and virtual environments in the study of the role of emotions in moral behavior. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), pp. 52–61. https://doi.org/10.1007/978-3-642-24571-8_6.

Panayiotou, G., Karekla, M., Georgiou, D., Constantinou, E., Paraskeva-Siamata, M., 2017a. Psychophysiological and self-reported reactivity associated with social anxiety and public speaking fear symptoms: effects of fear versus distress. Psychiatry Res. 255, 278–286. https://doi.org/10.1016/j.psychres.2017.05.044.

Panayiotou, G., Karekla, M., Georgiou, D., Constantinou, E., Paraskeva-Siamata, M., 2017b. Psychophysiological and self-reported reactivity associated with social anxiety and public speaking fear symptoms: effects of fear versus distress. Psychiatry Res. 255, 278–286. https://doi.org/10.1016/j.psychres.2017.05.044.

Panknin, T.L., Dickensheets, S.L., Nixon, S.J., Lovallo, W.R., 2002. Attenuated heart rate responses to public speaking in individuals with alcohol dependence. Alcohol. Clin. Exp. Res. 26, 841–847. https://doi.org/10.1111/J.1530-0277.2002.TB02613.X.

Parente, A.C.B.V., Garcia-Leal, C., Del-Ben, C.M., Guimarães, F.S., Graeff, F.G., 2005. Subjective and neurovegetative changes in healthy volunteers and panic patients performing simulated public speaking. Eur. Neuropsychopharmacol. 15, 663–671. https://doi.org/10.1016/j.euroneuro.2005.05.002.

Partala, T., Surakka, V., 2003. Pupil size variation as an indication of affective processing. Int. J. Hum. Comput. Stud. 59, 185–198. https://doi.org/10.1016/S1071-5819(03)00017-X.

Pereira, P., Duarte, E., Rebelo, F., Noriega, P., 2014. A review of gamification for health-related contexts. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). Springer Verlag, pp. 742–753. https://doi.org/10.1007/978-3-319-07626-3_70.

Pertaub, D.P., Slater, M., Barker, C., 2002. An experiment on public speaking anxiety in response to three different types of virtual audience. Presence Teleoperators Virtual Environ. 11, 68–78. https://doi.org/10.1162/105474602317343668.

Pertaub, D.P., Slater, M., Barker, C., 2001. An experiment on fear of public speaking in virtual reality. Studies in Health Technology and Informatics. SAGE Publications Inc., pp. 372–378. https://doi.org/10.3233/978-1-60750-925-7-372

Poeschl, S., 2017. Virtual reality training for public speaking-A QUEST-VR framework validation. Front. ICT 4, 13. https://doi.org/10.3389/FICT.2017.00013/BIBTEX.

Pörhölä, M., 2002. Arousal styles during public speaking. Commun. Educ. 51, 420–438. https://doi.org/10.1080/03634520216527.

Powers, M.B., Emmelkamp, P.M.G., 2008. Virtual reality exposure therapy for anxiety disorders: a meta-analysis. J. Anxiety Disord. 22, 561–569. https://doi.org/10.1016/J.JANXDIS.2007.04.006.

Pramana, G., Parmanto, B., Lomas, J., Lindhiem, O., Kendall, P.C., Silk, J., 2018. Using mobile health gamification to facilitate cognitive behavioral therapy skills practice in child anxiety treatment: open clinical trial. JMIR Serious Games 20, e8902. https://doi.org/10.2196/games.8902.

Premkumar, P., Heym, N., Brown, D.J., Battersby, S., Sumich, A., Huntington, B., Daly, R., Zysk, E., 2021. The effectiveness of self-guided virtual-reality exposure therapy for public-speaking anxiety. Front. Psychiatry 12, 1265. https://doi.org/10.3389/FPSYT.2021.694610/BIBTEX.

Pull, C.B., 2012. Current status of knowledge on public-speaking anxiety. Curr. Opin. Psychiatry. https://doi.org/10.1097/YCO.0b013e32834e06dc.

Raja, F.U., 2017. Anxiety level in students of public speaking: causes and remedies. J. Educ. Educ. Dev. 4, 94. https://doi.org/10.22555/joeed.v4i1.1001.

Reeves, R., Elliott, A., Curran, D., Dyer, K., Hanna, D., 2021. 360° Video virtual reality exposure therapy for public speaking anxiety: a randomized controlled trial. J. Anxiety Disord. 83, 102451 https://doi.org/10.1016/J.JANXDIS.2021.102451.

Riva, G., 2009. Virtual reality: an experiential tool for clinical psychology. Br. J. Guid. Couns. 37, 337–345. https://doi.org/10.1080/03069880902957056.

Riva, G., Davide, F., Ijsselsteijn, W., Riva, Giuseppe, 2003. Being There: concepts, effects and measurement of user presence in synthetic environments 1 Being There: the experience of presence in mediated environments.

Rothbaum, B.O., Hodges, L., Smith, S., Lee, J.H., Price, L., 2000. A controlled study of virtual reality exposure therapy for the fear of flying. J. Consult. Clin. Psychol. 68, 1020–1026. https://doi.org/10.1037/0022-006X.68.6.1020.

Ryu, J.H., Park, J.W., Nahm, F.S., Jeon, Y.T., Oh, A.Y., Lee, H.J., Kim, J.H., Han, S.H., 2018. The effect of gamification through a virtual reality on preoperative anxiety in pediatric patients undergoing general anesthesia: a prospective, randomized, and controlled trial. J. Clin. Med. 7, 284. https://doi.org/10.3390/jcm7090284.

Sailer, M., Homner, L., 2020. The gamification of learning: a meta-analysis. Educ. Psychol. Rev. 32, 77–112. https://doi.org/10.1007/s10648-019-09498-w.

Šalkevičius, J., Miškinyte, A., Navickas, L., 2019. Cloud based virtual reality exposure therapy service for public speaking anxiety. Inf 10, 62 10, 62. https://doi.org/10.3390/INFO10020062, 2019.

Schwerdtfeger, A., 2004. Predicting autonomic reactivity to public speaking: don't get fixed on self-report data! Int. J. Psychophysiol. 52, 217–224. https://doi.org/10.1016/J.IJPSYCHO.2003.10.008.

Skulmowski, A., Bunge, A., Kaspar, K., Pipa, G., 2014. Forced-choice decision-making in modified trolley dilemma situations: a virtual reality and eye tracking study. Front. Behav. Neurosci. 8 https://doi.org/10.3389/FNBEH.2014.00426/FULL.

Slater, M., 2002. Presence and the sixth sense. Presence Teleoperators Virtual Environ. https://doi.org/10.1162/105474602760204327.

Slater, M., Guger, C., Edlinger, G., Leeb, R., Pfurtscheller, G., Antley, A., Garau, M., Brogni, A., Friedman, D., 2006a. Analysis of physiological responses to a social situation in an immersive virtual environment. Presence: Teleoperators and Virtual Environments, pp. 553–569. https://doi.org/10.1162/pres.15.5.553.

Slater, M., Khanna, P., Mortensen, J., Yu, I., 2009a. Visual realism enhances realistic response in an immersive virtual environment. IEEE Comput. Graph. Appl. 29, 76–84. https://doi.org/10.1109/MCG.2009.55.

Slater, M., Lotto, B., Arnold, M.M., Sanchez-Vives, M.V., 2009b. How we experience immersive virtual environments: the concept of presence and its measurement. Anu. Psicol.

Slater, M., Pertaub, D.P., Barker, C., Clark, D.M., 2006b. An experimental study on fear of public speaking using a virtual environment. Cyberpsychol. Behav. 9, 627–633. https://doi.org/10.1089/CPB.2006.9.627.

Slater, M., Wilbur, S., 1997. A framework for immersive virtual environments (FIVE): speculations on the role of presence in virtual environments. Presence Teleoperators Virtual Environ. 6, 603–616. https://doi.org/10.1162/pres.1997.6.6.603.

Soash, R.L., 1999. Media equation: how people treat computers, television, and new media like real people and places. Collect. Manag. 24, 310–311. https://doi.org/10.1300/j105v24n03_14.

Spielberger, C.D., Gonzalez-Reigosa, F., Martinez-Urrutia, A., Natalicio, L.F.S., Natalicio, D.S., 1971. The state-trait anxiety inventory. Rev. Interam. Psicol. J. Psychol. 5, 3–4. https://doi.org/10.30849/RIP/IJP.V5I3.

Stein, M., Chavira, D., 1998a. Subtypes of social phobia and comorbidity with depression and other anxiety disorders. J. Affect. Disord. 50, S11–S16. https://doi.org/10.1016/S0165-0327(98)00092-5.

Stupar-Rutenfrans, S., Ketelaars, L.E.H., 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. Cyberpsychology, Behav. Soc. Netw. 20, 624–633. https://doi.org/10.1089/cyber.2017.0174.

Su, C.H., 2016. The effects of students' motivation, cognitive load and learning anxiety in gamification software engineering education: a structural equation modeling study. Multimed. Tools Appl. 75, 10013–10036. https://doi.org/10.1007/s11042-015-2799-7.

Takac, M., Collett, J., Blom, K.J., Conduit, R., Rehm, I., De Foe, A., 2019. Public speaking anxiety decreases within repeated virtual reality training sessions. PLoS ONE 14, e0216288. https://doi.org/10.1371/JOURNAL.PONE.0216288.

Teigen, K.H., 1994. Yerkes-dodson: a law for all seasons. Theory Psychol. 4, 525–547. https://doi.org/10.1177/0959354394044004.

Tinwell, A., 2014. The uncanny valley, in: the uncanny valley in games and animation. pp. 1–24. https://doi.org/10.1201/b17830-2.

Turpin, G., Grandfield, T., 2007. Electrodermal activity. Encyclopedia of Stress, pp. 899–902. https://doi.org/10.1016/B978-012373947-6.00139-2.

Vinayagamoorthy, V., Brogni, A., Gillies, M., Slater, M., Steed, A., 2004. An investigation of presence response across variations in visual realism. In: The 7th Annual International Presence Workshop, pp. 148–155.

Volante, M., Babu, S.V., Chaturvedi, H., Newsome, N., Ebrahimi, E., Roy, T., Daily, S.B., Fasolino, T., 2016. Effects of virtual human appearance fidelity on emotion contagion in affective inter-personal simulations. IEEE Trans. Vis. Comput. Graph. 22, 1326–1335. https://doi.org/10.1109/TVCG.2016.2518158.

Wakeford, N., 2002. The social life of avatars: presence and interaction in shared virtual environments. Sociol. Res. Online 7, 137–138. https://doi.org/10.1177/136078040200700211.

Westenberg, P.M., Bokhorst, C.L., Miers, A.C., Sumter, S.R., Kallen, V.L., van Pelt, J., Blöte, A.W., 2009. A prepared speech in front of a pre-recorded audience: subjective, physiological, and neuroendocrine responses to the leiden public speaking task. Biol. Psychol. 82, 116–124. https://doi.org/10.1016/J.BIOPSYCHO.2009.06.005.

Wiederhold, B.K., Riva, G., 2022. Metaverse creates new opportunities in healthcare. Annu. Rev. CyberTherapy Telemed.

Wienrich, C., Gross, R., Kretschmer, F., Muller-Plath, G., 2018. Developing and proving a framework for reaction time experiments in VR to objectively measure social interaction with virtual agents. In: 25th IEEE Conference on Virtual Reality and 3D User Interfaces, VR 2018 - Proceedings. Institute of Electrical and Electronics Engineers Inc., pp. 191–198. https://doi.org/10.1109/VR.2018.8446352

Yee, N., 2006. Motivations for play in online games. Cyberpsychology and Behavior. Mary Ann Liebert Inc., pp. 772–775. https://doi.org/10.1089/cpb.2006.9.772

Yoon, B., Kim II, H., Lee, G.A., Billinqhurst, M., Woo, W., 2019. The effect of avatar appearance on social presence in an augmented reality remote collaboration. In: 26th IEEE Conference on Virtual Reality and 3D User Interfaces, VR 2019 - Proceedings. Institute of Electrical and Electronics Engineers Inc., pp. 547–556. https://doi.org/10.1109/VR.2019.8797719

Zeiner, A.R., Smith, J.K., 1979. Cardiovascular and electrodermal predictors of SCR habituation among naive and experienced subjects. J. Clin. Psychol. 35, 510–512. https://doi.org/10.1002/1097-4679(197907)35:3<510::AID−JCLP2270350307>3.0.CO;2-X.

Zibrek, K., Cabral, J., McDonnell, R., 2021. Does synthetic voice alter social response to a photorealistic character in virtual reality?. In: Proceedings - MIG 2021: 14th ACM SIGGRAPH Conference on Motion, Interaction, and Games. Association for Computing Machinery, Inc. https://doi.org/10.1145/3487983.3488296.