Predicting the Effectiveness of Systematic Desensitization Through Virtual Reality for Mitigating Public Speaking Anxiety

Margaret von Ebers
Texas A&M University
mvonebers@tamu.edu

Ehsanul Haque Nirjhar
Texas A&M University
nirjhar71@tamu.edu

Amir H. Behzadan
Texas A&M University
abehzadan@tamu.edu

Theodora Chaspari
Texas A&M University
chaspari@tamu.edu

ABSTRACT

Public speaking is central to socialization in casual, professional, or academic settings. Yet, public speaking anxiety (PSA) is known to impact a considerable portion of the general population. This paper utilizes bio-behavioral indices captured from wearable devices to quantify the effectiveness of systematic exposure to virtual reality (VR) audiences for mitigating PSA. The effect of separate bio-behavioral features and demographic factors is studied, as well as the amount of necessary data from the VR sessions that can yield a reliable predictive model of the VR training effectiveness. Results indicate that acoustic and physiological reactivity during the VR exposure can reliably predict change in PSA before and after the training. With the addition of demographic features, both acoustic and physiological feature sets achieve improvements in performance. Finally, using bio-behavioral data from six to eight VR sessions can yield reliable prediction of PSA change. Findings of this study will enable researchers to better understand how bio-behavioral factors indicate improvements in PSA with VR training.

CCS CONCEPTS

• **Human-centered computing → Ubiquitous and mobile computing**.

KEYWORDS

Public speaking anxiety; virtual reality training; physiological signals; speech; wearable devices

ACM Reference Format:
Margaret von Ebers, Ehsanul Haque Nirjhar, Amir H. Behzadan, and Theodora Chaspari. 2020. Predicting the Effectiveness of Systematic Desensitization Through Virtual Reality for Mitigating Public Speaking Anxiety . In *Proceedings of the 2020 International Conference on Multimodal Interaction (ICMI '20), October 25–29, 2020, Virtual event, Netherlands.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3382507.3418883

1 INTRODUCTION

Speaking effectively to an audience is an essential skill in many facets of life, most importantly in the workplace. Yet, 15-30% of the population faces public speaking anxiety (PSA) [15], which impedes much of their academic and professional work and causes significant distress [34]. PSA is a communication-based disorder that involves physiological arousal, negative cognition, and behavioral reactions in response to a real or anticipated enactment of oral presentation [8]. Both self-reported and signal-based measures have been employed to quantify PSA. The former refers to the speakers' own views obtained through self-assessments, while the latter includes physiological responses of the autonomic nervous system (e.g., cardiovascular and electrodermal activity (EDA)), speech intonation, and visual information [1, 5, 11, 45]. Some individuals have different experiences of PSA than others. Female speakers tend to exhibit higher self-reported anxiety and physiological reactivity while speaking in public, compared to male speakers [6]. Non-native English speakers may have high fear of negative evaluation that can cause anxiety [26, 32, 39]. Prior works have also identified racial differences to PSA [13].

Several methods have been proposed for mitigating the negative effects of PSA. Cognitive restructuring attempts to teach individuals how to identify, evaluate, and modify negative feelings related to public speaking [3]. Frequent exposure to public speaking can yield the systematic desensitization of the stimuli and has the potential to gradually alter perceived negative thoughts in relation to public speaking [4]. Systematic exposure has been achieved with individuals presenting in front of a small-size real audience [23, 46], as well as through visual and virtual reality (VR) interfaces [5, 18, 19, 33]. Such interfaces rely on the provision of diverse naturalistic public speaking stimuli, rendering them potentially effective means to reducing PSA. The combination of VR exposure through multiple sessions (i.e. 4, 8) with cognitive restructuring has also been examined [2, 17, 43]. PSA treatments depict differential benefits for different groups of individuals. For example, cognitive restructuring was found to benefit more individuals with high social anxiety compared to their low anxiety counterparts [16, 38].

Previous studies in organizational and industrial psychology have attempted to predict the effect of computerized or guided job training modules using experiences from their past implementations in other locations [7, 20]. Employees' cognitive and affective reactions to such programs were found useful for predicting training success [41]. Yet, the literature in predicting the success of training interfaces is limited. To the best of our knowledge, it is not clear whether the effectiveness of VR training in mitigating PSA can be predicted from the individuals' reactivity, as well as the extent of training that is needed. In this paper, we examine the effectiveness of VR training in mitigating PSA through the following research questions:

- **RQ1**: Does VR training through systematic exposure to public speaking mitigate PSA?
- **RQ2**: Can we predict the effectiveness of the VR training by measuring individuals' bio-behavioral responses during the training?

ICMI '20, October 25–29, 2020, Virtual event, Netherlands

ACM ISBN 978-1-4503-7581-8/20/10…$15.00
https://doi.org/10.1145/3382507.3418883

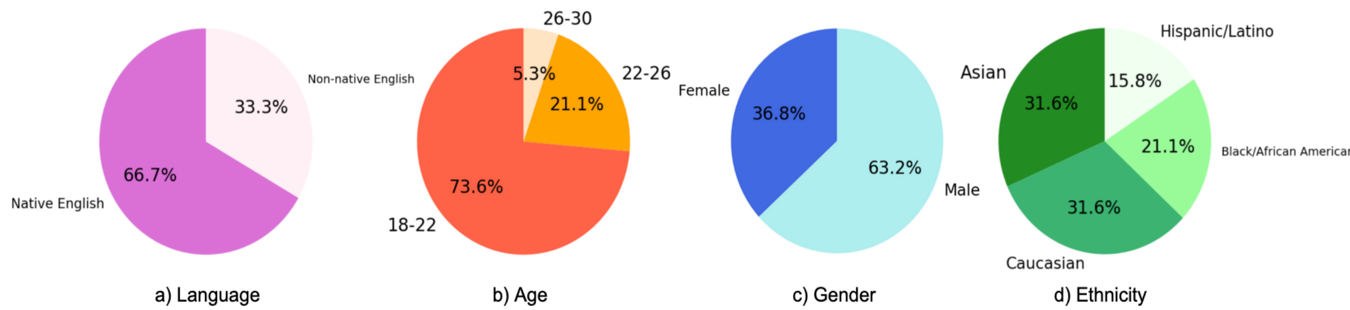


Figure 1: Distribution of participants based on demographic information.

- **RQ3**: Are there specific bio-behavioral indices and individual characteristics that are the most helpful in predicting changes in PSA occurring by VR training?
- **RQ4**: How many VR training sessions are needed to predict the effectiveness of the corresponding training?

To answer these questions, we conducted a research study that involved public speaking training with eight VR sessions. Physiological and acoustic indices were collected during the study, and pre-post differences before and after training were assessed through statistical analysis. Results demonstrated significant changes in PSA before and after VR training sessions. Bio-behavioral features can reliably estimate change in PSA with moderate to high Pearson's correlation of ($r = 0.49$, $p < 0.05$), when acoustic features are used. Addition of demographic factors further improves prediction performance. Finally, results indicate that data from six to eight VR training sessions is beneficial to the prediction of change in PSA.

2 PREVIOUS WORK

Prior works have focused on improving public speaking skills and mitigating PSA using diverse techniques. Batrinca *et al.* [5] and Chollet *et al.* [12] developed Cicero, a 2D avatar, that provides visual feedback to guide individuals through the public speaking task. Kimani *et al.* [22] proposed a virtual coach to promote the restructuring of negative thoughts during the public speaking encounter in the VR. Yadav *et al.* [45] and Sakib *et al.* [33] explored the effect of VR training for mitigating PSA using physiological and acoustic indices. These studies indicate that PSA can be reduced via systematic exposure to public speaking encounters, which can potentially lead to the desensitization of threatening stimuli [8]. Previous studies explored several ways to elicit PSA, including presenting pictures of social stimuli (e.g., faces) [14], instructing speech delivery to an imaginary audience [35], or presenting in front of a small-size real audience [23, 46]. VR can simulate a wide variety of naturalistic encounters that can be used for PSA exposure [27, 37]. VR experiences can elicit public speaking distress and repeated speaking sessions in VR can decrease anxiety levels [38, 40]. Harris *et al.* [17] used 4 such VR sessions for cognitive restructuring to reduce PSA, while Anderson *et al.* [2] and Wallach *et al.* [43] proposed cognitive behavioral therapy for PSA though 8 VR sessions. Despite promising results, the aforementioned studies have used statistical analysis to identify significant reduction in PSA without exploring the prediction of the PSA treatment effectiveness. Beyond PSA, previous studies in skill training have attempted to predict future performance as a function of training history [20, 41]. These works suggest the usefulness of modeling the process of skill acquisition and retention for predicting training effectiveness [21]. Predicting the potential benefit of a VR treatment to PSA might inform the design of effective training interfaces.

The contributions of this research to the existing body of knowledge are the following: (1) Previous works have explored models of state-based anxiety from bio-behavioral signals within a given presentation [11, 45], while this work aims to predict change in trait-based anxiety from the bio-behavioral data collected during VR training; (2) Previous works on VR training have identified a decrease in PSA after the end of the VR training through statistical analysis [19, 40]. Prediction of the potential benefit of VR training has not been performed; and (3) The extent to which individual characteristics moderate the association between PSA change and bio-behavioral measures has not been widely explored.

3 DATA DESCRIPTION

Our data comes from a laboratory user study with 23 participants (Fig. 1), who completed 10 public speaking sessions over a span of 2-3 weeks [45]. The study lasted 4 days per participant, where Days 1, 4 involved a single public speaking encounter with a real-life audience, and Days 2, 3 included the VR training (4 sessions per Day). Grounded on systematic desensitization treatments (Section 2), the public speaking training included 8 sessions in which participants had to present about news articles from topics of general interest in front of different types of VR stimuli. Participants were given 10 minutes to prepare for each presentation. Similar to prior works on VR training for PSA [13, 28, 33], 12 VR settings were generated based on different room conditions (i.e., board room, classroom, small theater, seminar room), audience reactions (i.e., negative, neutral, positive), and audience sizes (i.e., 12, 25, 54, 90). Each participant was randomly assigned to 8 of the 12 VR settings. Virtual environments were constructed through the Virtual Orator software [25] and displayed via the Oculus Rift headset [29]. Evaluation of PSA was performed in a pre-post experimental design, where each participant had a public speaking encounter with a real-life audience before and after the VR training.

Participants wore a wrist-mounted Empatica E4 [44] that collected EDA, blood volume pulse (BVP), body temperature, and 3-axis acceleration at sampling rates of 4, 64, 4, and 32 Hz, respectively. A Creative lavalier microphone was used to record presentations with a 16 kHz sampling rate and 16-bit recording. Demographic information of the participants were also collected. Demographic features included age, gender, ethnicity, and whether the corresponding participant was a native English speaker (Fig. 1). These will be referred to as "Age," "Gender," "Ethnicity," and "Language," respectively, in the rest of the paper. PSA was evaluated before and after the VR presentation through the Communication Anxiety Inventory (CAI) [10]. Participants completed CAI at the beginning of Days 1 and 4. The relative change in CAI before and after the VR training serves as an estimate of the the improvement in PSA. This assessment was chosen due to its relevance to PSA specifically.

4 METHODOLOGY

We will describe the pre-processing and feature extraction (Section 4.1), followed by the data analysis performed to answer our research questions (Sections 4.2).

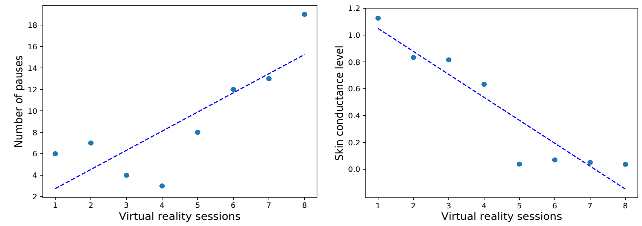
4.1 Data pre-processing and feature extraction

Consistent with prior works [24, 36, 45], pre-processing of the data included EDA outlier detection and replacement using linear interpolation, as well as voice activity detection (VAD) on audio signals to identify the presence and absence of speech. Feature extraction has been performed on both physiological and audio signals. Physiological features include 8 indices– mean skin conductance level (SCL), skin coductance response (SCR) amplitude and frequency, average heart rate, blood volume pulse (BVP), inter-beat interval (IBI), body temperature, and body acceleration. Acoustic features include 15 measures– the first 12 Mel-frequency cepstral coefficients (MFCC), jitter, shimmer, and number of pauses, since these have been found indicative of anxiety and emotional strain [11, 24, 42].

4.2 Data analysis

In order to answer RQ1, we perform a right-tailed paired t-test to identify significant differences in PSA before and after VR training. We answer RQ2 by performing machine learning experiments to predict the relative change in CAI before and after the VR training. Bio-behavioral features from VR sessions were used to fit a line for each participant, and the slope of this line was used as an input feature to predict change in CAI (Fig. 2). Regression models are evaluated through a 10-fold cross-validation. We note that each sample corresponds to a speaker, therefore there is no data contamination from the same speaker between training and test set. Pearson's correlation coefficient (r) between actual and estimated change in PSA is used as evaluation metric for the model.

We further explore RQ3 through statistical analysis and machine learning experiments. Statistical analysis includes a linear regression model to examine interactions between bio-behavioral features and demographic factors (i.e., Age, Gender, Language, and Ethnicity). Change in CAI is the dependent variable Y, while a bio-behavioral feature F, a demographic factor D, and their interaction are the independent variables of the model (i.e., $Y = c_1 \cdot F + c_2 \cdot D + c_3 \cdot F \cdot D$). Machine learning experiments are further



(a) Number of pauses (b) Skin conductance level

Figure 2: Examples of slopes across sessions used as features.

performed to estimated change in CAI in cross-validation framework, as in RQ2. The input features include different groups of acoustic and physiological features and their combination, in order to identify the usefulness of each group in the prediction task. We also add demographic factors in the feature space for conducting the regression task.

We finally answer RQ4 by performing another linear regression model to predict changes in PSA before and after the VR training. Instead of estimating the slope of bio-behavioral features from all eight sessions, as in RQ2 and RQ3, we used only features from first K sessions ($K = 2 \ldots, 8$). Evaluation of the regression models was performed as in RQ2. Results are visually inspected in order to identify temporal trajectories of the effectiveness of our approach in predicting PSA change using data from the first K training sessions.

5 RESULTS

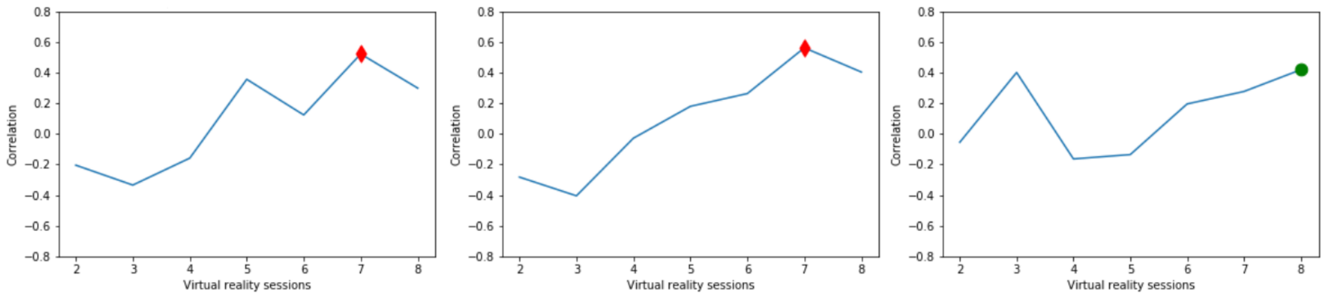
In response to RQ1, results indicate significant differences in CAI before and after the VR training ($t(22) = 2.64$, $p < 0.01$). It depicts that CAI scores among participants before training are significantly higher than CAI scores after training. For answering RQ2, the bio-behavioral features obtained from the eight VR sessions can successfully predict change in PSA in various cases (Table 1). Acoustic features appear to be useful in predicting the change of scores across VR training sessions with the combination of MFCCs and number of pauses performing the best ($r = 0.49$, $p < 0.05$). Physiological features demonstrated reduced predictive power ($r = 0.18$, $p = 0.45$). Each physiological feature was tested alone as well, and none had improved performance as compared to all physiological features combined. The slight difference with respect to the number of samples in the experiments stems from the fact all features were not available for all participants.

In response to RQ3, we examine the interaction effect between bio-behavioral features and demographic factors. We find that only Gender exhibited significant interaction effects with individual features (Table 2). Acoustic features, and specifically shimmer, depict

Table 1: Pearson's correlation coefficient (r) between actual and predicted change in public speaking anxiety using the slope of bio-behavioral features across the eight virtual reality sessions and individuals' demographic characteristics.

Bio-behavioral features	# Samples	Features only	Features & Age	Features & Language	Features & Ethnicity	Features & Gender
Physiological	19	0.18	0.20	0.55*	0.30	0.02
Jitter, shimmer	22	-0.16	0.25	-0.32	-0.19	-0.31
# Pauses	23	0.40	0.47*	-0.04	0.43*	0.31
MFCC	22	0.42*	0.35	-0.03	0.47*	0.37
MFCC, # pauses	23	0.49*	0.42*	-0.08	0.55**	0.46*
MFCC, jitter, shimmer	22	0.20	0.04	0.34	0.33	0.13
MFCC, jitter, shimmer, # pauses	23	0.47*	0.46*	0.19	0.43*	0.45*

*: $p < 0.05$, **: $p < 0.01$



a) Jitter, shimmer, and number of pauses

b) Number of pauses

c) Mel-frequency cepstral coefficients

Figure 3: Pearson's correlation between actual and predicted change in public speaking anxiety before and after the virtual reality (VR) training using linear regression that includes slopes of bio-behavioral features from K VR sessions ($K = 2, \ldots, 8$). The green circle and the red diamond markers indicate significant correlations ($p < 0.05$ and 0.01, respectively).

significant interaction effect with gender ($c_3 = -0.73$, $p = 0.02$). From the machine learning experiments (Table 1), the addition of demographic features can aid in the estimation of change in PSA. Age, Language, and Ethnicity were all found to best contribute to the correlations. Language information improved the correlation of the physiological features ($r = 0.55$, $p < 0.05$), but had no benefit on the acoustic. Age had a similar but reduced effect on physiological features ($r = 0.20$, $p = 0.42$) and depicted increased correlations with the number of pauses ($r = 0.47$, $p < 0.05$). Ethnicity was found to contribute to the feature set consisting of MFCCs and pauses ($r = 0.55$, $p < 0.01$), and slightly to the MFCCs ($r = 0.43$, $p < 0.05$) and pauses ($r = 0.47$, $p < 0.05$) alone.

Table 2: Regression coefficients obtained from linear regression with bio-behavioral features and gender.

Feature	Feature (c_1)	Gender (c_2)	Interaction (c_3)
SCR Frequency	-0.18	0.23	-0.01
Number of pauses	-0.97**	-0.33	0.87**
Shimmer	0.87**	0.38*	-0.73*

*: $p < 0.05$, **: $p < 0.01$

Finally for RQ4, we observe that the proposed models become more accurate in predicting change in PSA when including a large number of VR sessions (Fig. 3). A saturation is observed for the acoustic features, when reaching six to eight VR sessions (Figs. 3(a), (b)). It is also noteworthy that the MFCCs achieve good performance even when including the corresponding measures from only the first three sessions (Fig. 3(c)).

6 DISCUSSION

Our study revealed several interesting findings. Results indicate that individuals' bio-behavioral reactions obtained during the VR sessions can predict the effectiveness of the corresponding treatment in mitigating PSA, with acoustic measures being better indicators compared to physiological ones. Individual factors depict significant interactions with bio-behavioral features and can benefit the predictive ability of the machine learning models. Another potential reason that can explain the benefit of demographic factors is that many of the corresponding bio-behavioral measures are influenced by such factors. For example, acoustic features are heavily influenced by gender, age, and accent [9], therefore incorporating those factors in the feature space allows the models to learn the corresponding interactions. Table 2 exhibits significant interaction effects of gender and acoustic features. Inclusion of gender with the same features enhanced the predictive power of the model, as seen in Table 1. Finally, pauses alone were able to predict the effectiveness of VR training at a high degree. Prior work has highlighted the significance of pauses in capturing cognitive and affective processes in spontaneous speech [30, 31], which might be a potential reason for the high predictive ability of this feature.

The results reported in this paper should be considered in the light of the following limitations. Due to the longitudinal nature of the study, there are a relatively small number of participants, all of whom are either undergraduate or graduate students. This stage in life is often associated with public speaking in university classes and first job interviews, but additional study with an older user group may benefit adults who still experience public speaking anxiety. Additionally, the study discussed the effects of eight training sessions within a short span, which might introduce habituation effects. As part of future work, participants may complete more sessions over a larger period of time in order to reduce habituation and study how much training is needed. Finally, a control condition was not examined in this study. Comparing the VR training to another type of training (e.g., real-life audience) might afford us additional insights into the specific cognitive and psychological processes that are attributed to the VR environment.

7 CONCLUSION

This paper examined quantifiable estimators of change in PSA before and after VR training and the extent to each repeated VR sessions to alleviate PSA. Results identified specific sets of acoustic and physiological features that can predict change in PSA, with demographic factors substantially improving the predictive ability of the model. Effective prediction of the PSA change can occur by including data from six to eight VR sessions. The results create new opportunities for future work, such as determining if the reduction in anxiety lasts over long periods of time. Future research may also involve improving on the limitations of this study, including the collection of a larger sample size and repeating the study with groups beyond university students.

ACKNOWLEDGMENTS

The authors would like to acknowledge the Engineering Information Foundation (EiF) for funding this work (grant EiF 18.02).

REFERENCES

[1] Ana C Aguiar, Mariana Kaiseler, Hugo Meinedo, Traian E Abrudan, and Pedro R Almeida. 2013. Speech stress assessment using physiological and psychological measures. In *Proceedings of the 2013 ACM conference on Pervasive and ubiquitous computing adjunct publication*. 921–930.

[2] Page L Anderson, Matthew Price, Shannan M Edwards, Mayowa A Obasaju, Stefan K Schmertz, Elana Zimand, and Martha R Calamaras. 2013. Virtual reality exposure therapy for social anxiety disorder: A randomized controlled trial. *Journal of consulting and clinical psychology* 81, 5 (2013), 751.

[3] Page L Anderson, Elana Zimand, Larry F Hodges, and Barbara O Rothbaum. 2005. Cognitive behavioral therapy for public-speaking anxiety using virtual reality for exposure. *Depression and anxiety* 22, 3 (2005), 156–158.

[4] Joe Ayres and Tim Hopf. 1992. Visualization: Reducing speech anxiety and enhancing performance. *Communication Reports* 5, 1 (1992), 1–10.

[5] Ligia Batrinca, Giota Stratou, Ari Shapiro, Louis-Philippe Morency, and Stefan Scherer. 2013. Cicero-towards a multimodal virtual audience platform for public speaking training. In *International workshop on intelligent virtual agents*. Springer, 116–128.

[6] Ralph R Behnke and Chris R Sawyer. 2000. Anticipatory anxiety patterns for male and female public speakers. *Communication education* 49, 2 (2000), 187–195.

[7] Bianca Bier, Émilie Ouellet, and Sylvie Belleville. 2018. Computerized attentional training and transfer with virtual reality: Effect of age and training type. *Neuropsychology* 32, 5 (2018), 597.

[8] Graham D Bodie. 2010. A racing heart, rattling knees, and ruminative thoughts: Defining, explaining, and treating public speaking anxiety. *Communication education* 59, 1 (2010), 70–105.

[9] Kamonpan Boonkit. 2010. Enhancing the development of speaking skills for non-native speakers of English. *Procedia-social and behavioral sciences* 2, 2 (2010), 1305–1309.

[10] Steven Booth-Butterfield and Malloy Gould. 1986. The communication anxiety inventory: Validation of state-and context-communication apprehension. *Communication Quarterly* 34, 2 (1986), 194–205.

[11] Lei Chen, Gary Feng, Jilliam Joe, Chee Wee Leong, Christopher Kitchen, and Chong Min Lee. 2014. Towards automated assessment of public speaking skills using multimodal cues. In *Proc. ICMI*. ACM, 200–203.

[12] Mathieu Chollet, Torsten Wörtwein, Louis-Philippe Morency, Ari Shapiro, and Stefan Scherer. 2015. Exploring feedback strategies to improve public speaking: An interactive virtual audience framework. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 1143–1154.

[13] Julia Diemer, Andreas Mühlberger, Paul Pauli, and Peter Zwanzger. 2014. Virtual reality exposure in anxiety disorders: impact on psychophysiological reactivity. *The World Journal of Biological Psychiatry* 15, 6 (2014), 427–442.

[14] Ulf Dimberg, Mats Fredrikson, and Olle Lundquist. 1986. Autonomic reactions to social and neutral stimuli in subjects high and low in public speaking fear. *Biological Psychology* 23, 3 (1986), 223–233.

[15] National Collaborating Centre for Mental Health et al. 2013. *Social anxiety disorder: the NICE guideline on recognition, assessment and treatment*. Royal College of Psychiatrists.

[16] William J Fremouw and Robert E Zitter. 1978. A comparison of skills training and cognitive restructuring-relaxation for the treatment of speech anxiety. *Behavior Therapy* 9, 2 (1978), 248–259.

[17] Sandra R Harris, Robert L Kemmerling, and Max M North. 2002. Brief virtual reality therapy for public speaking anxiety. *Cyberpsychology & behavior* 5, 6 (2002), 543–550.

[18] Brian L Heuett and Kyle B Heuett. 2011. Virtual reality therapy: A means of reducing public speaking anxiety. *International Journal of Humanities and Social Science* 1, 16 (2011), 1–6.

[19] Francisco-Javier Hinojo-Lucena, Inmaculada Aznar-Díaz, María-Pilar Cáceres-Reche, Juan-Manuel Trujillo-Torres, and José-María Romero-Rodríguez. 2020. Virtual Reality Treatment for Public Speaking Anxiety in Students. Advancements and Results in Personalized Medicine. *Journal of Personalized Medicine* 10, 1 (2020), 14.

[20] V Joseph Hotz, Guido W Imbens, and Julie H Mortimer. 2005. Predicting the efficacy of future training programs using past experiences at other locations. *Journal of Econometrics* 125, 1-2 (2005), 241–270.

[21] Matt Jones, Lyle E Bourne Jr, and Alice F Healy. 2012. A compact mathematical model for predicting the effectiveness of training. (2012).

[22] Everlyne Kimani and Timothy Bickmore. 2019. Addressing Public Speaking Anxiety in Real-time Using a Virtual Public Speaking Coach and Physiological Sensors. In *Proceedings of the 19th ACM International Conference on Intelligent Virtual Agents*. 260–263.

[23] Clemens Kirschbaum, Jens C Prussner, Arthur A Stone, Ilona Federenko, Jens Gaab, Doris Lintz, Nicole Schommer, and Dirk H Hellhammer. 1995. Persistent high cortisol responses to repeated psychological stress in a subpopulation of healthy men. *Psychosomatic medicine* 57, 5 (1995), 468–474.

[24] Ehsanul Haque Nirjhar, Amir Behzadan, and Theodora Chaspari. 2020. Exploring Bio-Behavioral Signal Trajectories of State Anxiety During Public Speaking. In *ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 1294–1298.

[25] Virtual Orator. 2020. https://virtualorator.com/

[26] Eylem Oruç and Cavide Demirci. 2020. Foreign language anxiety and English language achievement in higher education: The mediating role of student engagement. *European Journal of Education Studies* (2020).

[27] DP Pertaub, M Slater, and C Barker. 2001. An experiment on fear of public speaking in virtual reality. *Studies in health technology and informatics* (2001), 372–378.

[28] David Paul Pertaub, Mel Slater, and Chris Barker. 2002. An experiment on public speaking anxiety in response to three different types of virtual audience. *Presence: Teleoperators & Virtual Environments* 11, 1 (2002), 68–78.

[29] Oculus Rift. 2020. https://www.oculus.com/

[30] Sherry R Rochester. 1973. The significance of pauses in spontaneous speech. *Journal of Psycholinguistic Research* 2, 1 (1973), 51–81.

[31] Ralph Leon Rose. 1998. The communicative value of filled pauses in spontaneous speech. *MA Diss., Univ. of Birmingham* (1998).

[32] Firooz Sadighi and Mehdi Dastpak. 2017. The sources of foreign language speaking anxiety of Iranian English language learners. *International Journal of Education and Literacy Studies* 5, 4 (2017), 111–115.

[33] Md. Nazmus Sakib, Megha Yadav, Theodora Chaspari, and Amir H. Behzadan. 2019. Coupling virtual reality and physiological markers to improve public speaking performance. In *International Conference on Construction Applications of Virtual Reality*.

[34] Franklin R Schneier. 2006. Social anxiety disorder. *New England Journal of Medicine* 355, 10 (2006), 1029–1036.

[35] Andreas Schwerdtfeger. 2004. Predicting autonomic reactivity to public speaking: don't get fixed on self-report data! *International Journal of Psychophysiology* 52, 3 (2004), 217–224.

[36] Lin Shu, Jinyan Xie, Mingyue Yang, Ziyi Li, Zhenqi Li, Dan Liao, Xiangmin Xu, and Xinyi Yang. 2018. A review of emotion recognition using physiological signals. *Sensors* 18, 7 (2018), 2074.

[37] Mel Slater, D-P Pertaub, and Anthony Steed. 1999. Public speaking in virtual reality: Facing an audience of avatars. *IEEE Computer Graphics and Applications* 19, 2 (1999), 6–9.

[38] Snežana Stupar-Rutenfrans, Loes EH Ketelaars, and Marnix S van Gisbergen. 2017. Beat the fear of public speaking: Mobile 360 video virtual reality exposure training in home environment reduces public speaking anxiety. *Cyberpsychology, Behavior, and Social Networking* 20, 10 (2017), 624–633.

[39] Gonca Subaşı. 2010. What are the main sources of Turkish EFL students' anxiety in oral practice? *Turkish Online Journal of Qualitative Inquiry* 1, 2 (2010), 29–49.

[40] Marcel Takac, James Collett, Kristopher J Blom, Russell Conduit, Imogen Rehm, and Alexander De Foe. 2019. Public speaking anxiety decreases within repeated virtual reality training sessions. *PloS one* 14, 5 (2019).

[41] James A Tan, Rosalie J Hall, and Carol Boyce. 2003. The role of employee reactions in predicting training effectiveness. *Human Resource Development Quarterly* 14, 4 (2003), 397–411.

[42] Dimitrios Ververidis and Constantine Kotropoulos. 2006. Emotional speech recognition: Resources, features, and methods. *Speech communication* 48, 9 (2006), 1162–1181.

[43] Helene S Wallach, Marilyn P Safir, and Margalit Bar-Zvi. 2009. Virtual reality cognitive behavior therapy for public speaking anxiety: a randomized clinical trial. *Behavior modification* 33, 3 (2009), 314–338.

[44] Empatica E4 Wristband. 2020. https://www.empatica.com/

[45] Megha Yadav, Md Nazmus Sakib, Kexin Feng, Theodora Chaspari, and Amir Behzadan. 2019. Virtual reality interfaces and population-specific models to mitigate public speaking anxiety. In *2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII)*. IEEE, 1–7.

[46] Antonio Waldo Zuardi, José Alexandre de Souza Crippa, Jaime Eduardo Cecílio Hallak, and Ricardo Gorayeb. 2013. Human experimental anxiety: actual public speaking induces more intense physiological responses than simulated public speaking. *Brazilian Journal of Psychiatry* 35, 3 (2013), 248–253.