Affective and Pleasurable Design, Vol. 204, 2026, 167–174
https://doi.org/10.54941/ahfe1007342

Effects of Virtual Reality-Based Speech Practice on Psychological States and Performance

Itsuki Kubota[1], Kazunori Kaede[1,2], and Keiichi Watanuki[1,2]

[1]Graduate School of Science and Engineering, Saitama University, 255 Shimo-okubo, Sakura-ku, Saitama-shi, Saitama 338-8570, Japan
[2]Advanced Institute of Innovative Technology, Saitama University, 255 Shimo-okubo, Sakura-ku, Saitama-shi, Saitama 338-8570, Japan

ABSTRACT

Many individuals experience anxiety while delivering a speech. One therapeutic approach for mitigating such anxiety is exposure therapy, in which individuals are gradually exposed to anxiety-inducing situations of increasing intensity. A psychological intervention known as Virtual Reality Exposure Therapy (VRET) has been developed in recent years. VRET enables simulated exposure therapy by constructing anxiety-provoking environments within a virtual space and having the participants wear a head-mounted display. Prior studies targeting speech anxiety have used virtual reality (VR) speech venues in which the audience's attitude and size were manipulated to vary the level of perceived anxiety. This study aimed to determine the most effective repetition pattern of practice and to propose a more effective method for VR-based speech training. In the experiment, we established three anxiety levels for the VR speech venue. Participants then participated in VR speech practice sessions in which their anxiety levels were manipulated. The participants were divided into four groups: groups that repeatedly practiced at Level 1, Level 2, and Level 3 and a group that practiced with gradually increasing levels from 1 to 3. Before and after the VR practice sessions, the participants delivered face-to-face speeches to evaluate changes in performance. For face-to-face speeches, both self-evaluations by the speakers and external evaluations by audience members (experimental assistants) were collected. In this experiment, changes in the pattern of practice repetition did not have a substantial effect on the psychological state or performance. As opposed to the specific pattern of repetition, repeated practice itself and the resulting habituation may contribute to reductions in nervousness and anxiety as well as improvements in subjective performance evaluations.

Keywords: Public speaking anxiety, Virtual reality, Performance, Speech audience, Physiological monitoring

INTRODUCTION

Speech anxiety is common. A previous study reported that 60% of university students surveyed experienced anxiety about public speaking, and that such anxiety may lead to lower perceived interpersonal evaluation and self-efficacy (Marinho et al., 2017). Exposure therapy is used to reduce various forms of anxiety, including speech anxiety. Gradual exposure is commonly employed,

Received April 15, 2026; Revised May 31, 2026; Accepted June 18, 2026; Available online July 20, 2026

in which the level of anxiety-inducing situations is increased step by step to promote habituation to anxiety and thereby reduce it.

A method known as Virtual Reality Exposure Therapy (VRET) has been developed in recent years. In VRET, users wear a head-mounted display (HMD) and can easily perform simulated exposure therapy at home in a virtual environment. A previous study on VRET for speech anxiety conducted repeated virtual reality (VR) speech sessions while gradually increasing the anxiety levels in a virtual speech venue (Premkumar et al., 2021). The results indicated a decrease in both subjective anxiety and heart rate. However, as with conventional exposure therapy, few studies have clearly demonstrated that gradual exposure, that is, stepwise increases in anxiety levels, is the most effective approach for VRET.

The present study investigated how different methods of increasing and repeating anxiety levels in VR speech tasks affect the reduction of speakers' anxiety and nervousness, as well as improvements in their performance. To evaluate the performance, face-to-face speeches were conducted in front of a real audience before and after the VR speech sessions, and both self-evaluations and evaluations by others were used to obtain subjective and objective assessments. In addition, physiological data were measured throughout all speech sessions to clarify changes in anxiety and nervousness from a physiological perspective.

Experimental Procedure

The experiment was conducted with 12 participants (mean age: 21.5 ± 1.3 years). This study was approved by the Ethics Committee for Human Subject Research of Saitama University (Approval No. R7-E-28). Informed consent was obtained from all participants prior to the experiment.

Physiological measurements were recorded during both the rest period and speech tasks. During the VR speech sessions, the participants wore an HMD (Meta Quest 2, manufactured by Meta Platforms). A multichannel physiological measurement device (NeXus10 MARK II, manufactured by Mind Media) was used to measure the blood volume pulse (LF/HF) and skin conductance (SC).

During the rest period, participants maintained a relaxed state while looking at a fixation cross. The preparation period was used for thinking about the contents of the speech. One set of each speech task consisted of a 5-min rest period, 1-min preparation period, 5-min speech, and questionnaire.

The procedure was conducted as follows: a topic selection questionnaire, the first face-to-face speech session (Face-to-Face Speech 1), six VR speech sessions, and the second face-to-face speech session (Face-to-Face Speech 2). Each session was conducted on a different day.

Topic Selection Questionnaire

Conditions regarding the speech topics were established, and participants completed a topic selection questionnaire. Participants subjectively rated how much anxiety and nervousness they would feel if they were asked to

deliver a speech on each of the prepared topics (e.g., favorite foods and drinks, extracurricular activities they had belonged to, and edible insects).

The questionnaire consisted of 40 topics designed from different perspectives, such as whether the topic was personally relevant and familiar to the participant. Their responses were reflected in the settings of the VR speech sessions.

Face-To-Face Speech

In the face-to-face speech sessions, participants delivered speeches in front of a single audience member in a real setting. A different person served as the audience member in each session and was instructed not to react. The speaker and audience member sat on chairs placed 1.5 m apart.

After each speech, the speaker completed a questionnaire regarding anxiety and nervousness, as well as a performance questionnaire for self-evaluation. The audience member completed a performance questionnaire for the other evaluation.

VR SPEECH

Three anxiety levels were defined for the VR speech sessions: low, medium, and high. The 12 participants were divided into four groups of three participants each, and the manner in which the anxiety levels were increased and repeated over the six VR speech sessions differed among the groups. The patterns of anxiety-level progression in Groups 1–4 (G1, G2, G3, and G4) are shown in Table 1. In G1, the anxiety level was increased gradually, whereas in G2, G3, and G4, the low, medium, and high levels, respectively, were repeated throughout the six sessions.

Anxiety levels were defined by combining the conditions related to the video environment, reward system, and speech topics. In terms the video environment, the audience's attitude and crowding levels varied across the videos. Three videos were prepared, as follows: Video 1 (low anxiety level), Video 2 (medium anxiety level), and Video 3 (high anxiety level). The contents of each video are shown in Table 2, and Figure 1 presents an example of Video 3. The videos were recorded using a 360-degree camera and presented on the HMD.

In addition, reward conditions were introduced. Before the speech, the participants were given snacks. Depending on the score of the audience member's performance evaluation, the snacks were either retained or removed. This condition was intended to increase the participants' nervousness and to encourage greater concentration during the speech.

By combining these conditions, the following anxiety levels were defined for all participants in the VR speech sessions:

Low anxiety level: Video 1, no reward condition, and a topic that was unlikely to evoke anxiety.

Medium anxiety level: Video 2, reward condition present, and a topic that evoked a moderate level of anxiety.

High anxiety level: Video 3, reward condition present, and a topic that evoked a high level of anxiety.

Table 1: Description of the video conditions.

Group	Exposure Type	Repetition Pattern					
		1st	2nd	3rd	4th	5th	6th
1	Gradual	Low	Low	Mid	Mid	High	High
2	Repeated low level	Low	Low	Low	Low	Low	Low
3	Repeated medium level	Mid	Mid	Mid	Mid	Mid	Mid
4	Repeated high level	High	High	High	High	High	High

Table 2: VR speech grouping and repetition.

Video	Audience Size	Audience Attitude	Venue Size (m^2)
1	0	-	10
2	3	Neutral, no facial expression, no reaction	20
3	20	Negative, rejective	40

Figure 1: VR speech auditorium.

Evaluation Indices

LF/HF is an index of autonomic nervous system activity that is used as an indicator of stress. SC, which measures mental sweating, is used to evaluate nervousness and relaxation. Higher values of LF/HF and SC indicate greater dominance of the sympathetic nervous system. LF/HF and SC were calculated during both the pre-rest and speech periods.

The questionnaire regarding anxiety and nervousness consisted of eight items rated on a seven-point scale, with higher scores indicating stronger feelings of anxiety or nervousness.

The performance questionnaire consisted of seven items rated on a seven-point scale, and two versions were prepared: one for self-evaluation and one for evaluation by others. Higher scores indicated better performance, and the questionnaire addressed aspects such as speech organization and style.

Experimental Results

The experiment was divided into the following phases: Face-to-Face Speech 1, Early VR (the 1st and 2nd VR speech sessions), Middle VR (the 3rd and 4th sessions), Late VR (the 5th and 6th sessions), and Face-to-Face Speech 2. Statistical analyses were conducted between phases and between groups at a significance level of 5%.

For LF/HF and SC, the differences between the values during the speech period and the rest period were calculated for each condition, and statistical analyses were performed on these values as ΔLF/HF and ΔSC.

A mixed-design analysis of variance (ANOVA) was conducted on ΔLF/HF. The main effect of group was significant, whereas the main effect of phase and the interaction between phase and group were not significant. The results of ΔLF/HF across groups are shown in Figure 2. To examine the main effect of group, t-tests with Holm correction were performed using the mean ΔLF/HF across all phases for each participant. A significant difference was found between G1 and G2.

A mixed-design ANOVA was also conducted on ΔSC. The main effect of phase was significant, whereas the main effect of group and the interaction between phase and group were not significant. The results of ΔSC across phases are shown in Figure 3. To examine the main effect of phase, t-tests with Holm correction were conducted. The Early VR phase showed significantly higher values than both the Middle VR and Late VR phases. In addition, the Middle VR phase showed significantly higher values than the Late VR phase.

For the questionnaires, the mean score of the questionnaire items at each measurement session was calculated and used as a representative value for each participant. Statistical analyses were then conducted on these representative values across phases and groups.

The results of the anxiety and nervousness questionnaire for each group are shown in Figures 4(a)–(d). A Friedman test was conducted to compare the phases, and a significant overall trend was observed. However, when a Wilcoxon signed-rank test was performed, no significant differences were found between any specific phases.

To compare the groups, a Kruskal–Wallis test was conducted. A trend toward a significant difference among the groups was found in the Late VR phase. However, when a Mann–Whitney U test with Bonferroni correction was performed, no significant differences were found between any specific groups.

For the two performance questionnaires (self-evaluation and other-evaluation), the results comparing Face-to-Face Speech 1 and Face-to-Face Speech 2 within each group are shown in Figures 5 and 6. A Wilcoxon signed-rank test was conducted to compare Face-to-Face Speech 1 and Face-to-Face Speech 2 within each group, and no significant differences were found in either self-evaluation or other-evaluation.

Although no significant difference was observed in self-evaluation, all groups showed higher scores in Face-to-Face Speech 2 than in Face-to-Face Speech 1.

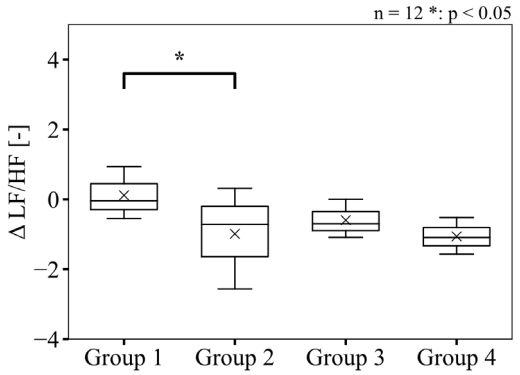

Figure 2: Group differences in ΔLF/HF.

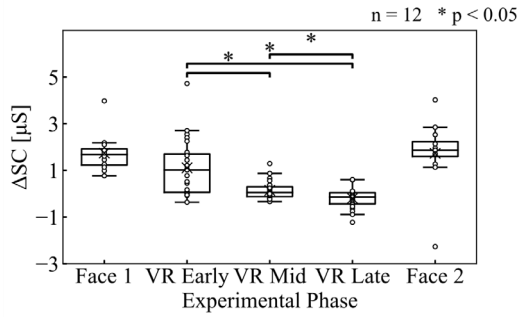

Figure 3: Between phase comparison of ΔSC.

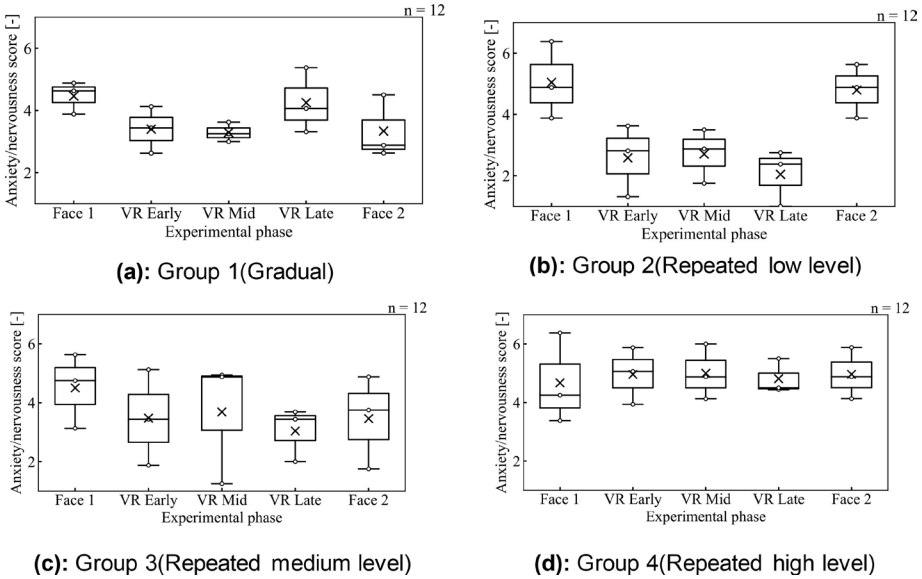

Figure 4: Anxiety tension score for each gruop.

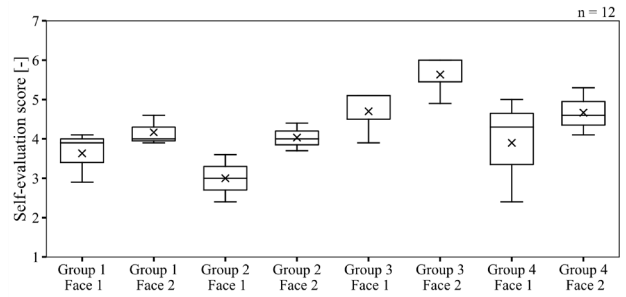

Figure 5: Self-evaluation score.

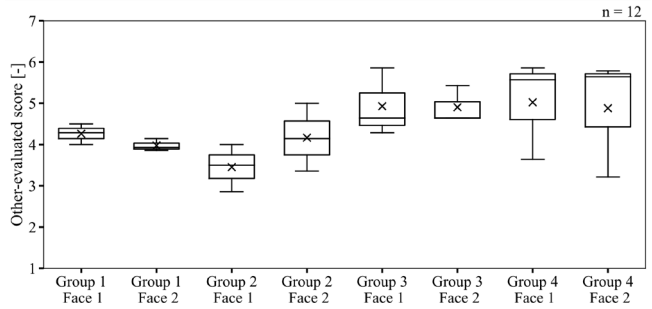

Figure 6: Other evaluation score.

DISUCUSSION

As shown in Figure 2, a significant difference in ΔLF/HF was observed only between G1 and G2. G2 was the group in which the low-anxiety condition was repeated, and it is possible that excessive stress responses were less likely to occur throughout the experiment, resulting in greater physiological stability.

In contrast, as shown in Figure 3, ΔSC tended to change across the phases, but no significant differences were observed among the groups. This suggests that ΔSC was influenced more strongly by the repetition of the VR speech task than by differences in anxiety levels. ΔSC gradually decreased from the Early VR phase to the Middle and Late VR phases, indicating that physiological arousal in response to the speech task gradually decreased and that participants became accustomed to the speech task.

Therefore, ΔLF/HF may be more susceptible to differences between groups, whereas ΔSC may be more susceptible to differences between phases. However, the interaction between phase and group was not significant for either ΔLF/HF or ΔSC. Thus, it could not be confirmed whether differences in the manner in which the anxiety levels were increased or repeated altered the pattern of change across phases or affected changes in physiological responses.

In terms of the questionnaire results, Figure 4 shows that the anxiety and nervousness questionnaires exhibited an overall trend across the phases, although no significant differences were detected between any specific phases.

This suggests that even when VR speech sessions such as those in the present experiment are repeated, reductions in subjective anxiety and nervousness may be difficult to observe.

In addition, Figure 5 shows that the self-evaluated performance tended to be higher in Face-to-Face Speech 2 than in Face-to-Face Speech 1. This suggests that subjective performance evaluations may be improved through repeated VR speech sessions.

Figure 6 shows that no significant difference was observed in the other-evaluated performances between Face-to-Face Speech 1 and Face-to-Face Speech 2. This suggests that clear changes in performance may not be perceived by the audience in a short-term VR speech intervention.

CONCLUSION

This study investigated the effects of different repetition patterns in VR speech practice on physiological responses, anxiety, nervousness, and performance. The results of the experiment suggest that ΔLF/HF may be more sensitive to differences between groups, whereas ΔSC may be more sensitive to changes across phases. However, changes in the pattern of practice repetition did not have a substantial effect on the psychological state or performance.

As opposed to the specific pattern of repetition, repeated practice itself and the resulting habituation may contribute to reductions in nervousness and anxiety as well as improvements in subjective performance evaluations. Future studies should conduct longer-term experiments to examine long-term changes in the psychological state and performance.

REFERENCES

Marinho, A. C. F., de Medeiros, A. M., Gama, A. C. C., Teixeira, L. C. (2017) "Fear of Public Speaking: Perception of College Students and Correlates," *Journal of Voice*, Vol. 31, No. 1, pp. 127.e7–127.e11.

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," *Frontiers in Psychiatry*, Vol. 12.